Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Pure and Applied Algebra 200 (2005) 163–190
JOURNAL OF PURE AND APPLIED ALGEBRA
www.elsevier.com/locate/jpaa

Differential algebra for derivations with nontrivial commutation rules

Evelyne Hubert

INRIA Sophia Antipolis, Sophia Antipolis 06902, France

Received 21 October 2003; received in revised form 20 October 2004
Available online 2 March 2005
Communicated by M.-F. Roy

Abstract

The classical assumption of differential algebra, differential elimination theory and formal integrability theory is that the derivations do commute. This is the standard case arising from systems of partial differential equations written in terms of the derivations w.r.t. the independant variables. We inspect here the case where the derivations satisfy nontrivial commutation rules. Such a situation arises, for instance, when we consider a system of equations on the differential invariants of a Lie group action. We develop the algebraic foundations for such a situation. They lead to algorithms for completion to formal integrability and differential elimination.

MSC: 12H05, 12H20, 53A55, 16W22, 16W25, 16W99, 16S15, 16S30, 16S32

1. Introduction

We establish the bases of a differential algebra theory, aimed at differential elimination, where the derivations do not commute but satisfy some non trivial relationships.

Classically [21,36], to treat algebraic differential systems with independent variables $(t_1, \ldots, t_m)$ and dependent variables $\mathscr{Y} = \{y_1, \ldots, y_n\}$ we introduce the ring of differential polynomials $\mathscr{F}[y_\alpha \mid y \in \mathscr{Y}, \alpha \in \mathbb{N}^m]$ where $\mathscr{F}$ is a field of rational or meromorphic functions in $(t_1, \ldots, t_m)$. $\mathscr{F}$ is naturally endowed with the commuting derivations

E-mail address: evelyne.hubert@inria.fr.
URL: http://www-sop.inria.fr/cafe/evelyne.hubert.

0022-4049/$ - see front matter
doi:10.1016/j.jpaa.2004.12.034

$\delta_1 = \partial/\partial t_1, \ldots, \delta_m = \partial/\partial t_m$. These derivations are extended to $\mathscr{F}[y_\alpha \mid y \in \mathscr{Y}, \alpha \in \mathbb{N}^m]$ by the formula $\delta_k(y_\alpha) = y_{\alpha+\varepsilon_k}$ where ε_k is the m-uplet having 1 as kth component and 0 otherwise. We call $\mathscr{Y}$ the set of differential indeterminates while $\{y_\alpha \mid y \in \mathscr{Y}, \alpha \in \mathbb{N}^m\}$ is the set of their derivatives. To make the link with differential geometry, $\mathscr{F}[y_\alpha \mid y \in \mathscr{Y}, \alpha \in \mathbb{N}^m]$ is the coordinate ring of the infinite jet space and the δ_i are the total derivatives w.r.t. the independent variables.

In this paper, we treat differential systems in some differential indeterminates $\mathscr{Y} = \{y_1, \ldots, y_n\}$ with derivations $\{\delta_1, \ldots, \delta_m\}$ that do not commute but rather satisfy commutation rules of the type

$$\delta_i \delta_j = \delta_i \delta_j + \sum_{l=1}^{m} c_{ijl} \delta_l,$$

where the coefficients c_{ijl} are polynomials[1] in the derivatives of $\mathscr{Y}$.

The first difficulty here is to define the differential polynomial ring. In classical differential algebra the derivations are defined on the coefficient field $\mathscr{F}$ and extended to the polynomial differential ring. Indeterminates can be introduced one by one. A first extension consists in considering (finite sets) the derivations on $\mathscr{F}$ that satisfy nontrivial commutation rules [22,40]. The considered derivations generate a subspace of the $\mathscr{F}$-vector space of derivations on $\mathscr{F}$. An algebraic version of Frobenius theorem [11, Proposition 39; 22, Section 0.5, Proposition 6] shows that you can always choose a commuting basis of derivations for that vector subspace provided it is closed under the commutator. There are nonetheless theoretical and computational advantages to work with non commuting derivations in this case.

In the case we consider, the commutators of the derivations are to be equal to a linear combination of the derivations but the coefficients belong to the differential polynomial ring to be defined. The field of coefficients is actually a field of constants for the derivations. It is therefore not possible to attack the problem by an analogue of the Poincare, de Witt and Birkhoff theorem, i.e. by exhibiting a normal form for the derivation operators of order 2 or more. We need to show a normal form for the differential polynomials directly.

Our motivation for this generalization of differential algebra takes its root in a project initiated by Mansfield. One has to reckon that differential systems that are *too* symmetric lead to intractable computations for differential elimination software, as for instance [5,29,39]. The introduction of a ranking in the underlying algorithms indeed breaks the symmetry instead of using it to reduce the problem. Mansfield's original idea was to factor systems invariant under the action of a Lie group by their symmetry before their treatment. The moving frame construction proposed by Fels and Olver [13] provides the ingredients of a reduction by the symmetry. A first reduction was proposed in [30]. The present paper offers the differential algebra foundations for a different reduction. We shall give a quick example of this reduction as a motivating example. The complete description of the reduction used is presented and compared to the reduction of [30] in [18].

[1] For the general case where the c_{ijl} are rational functions we can mimic a localization by introducing a new indeterminate that represents the inverse of the common denominator. There is thus no loss of generality in considering only polynomials as coefficients.

Actually many a differential elimination problems draw advantages from being expressed in terms of non standard and often non commuting derivations. That is the case of equivalence problems [12,33] and classification problems [25,26]. We provide the algebraic ground for this. For a wider applicability we start from the commutation rules, contrary to [23] that is based on a change of derivations.

Before we sketch the plan of the paper, let us mention that several recent works consider algebras with nontrivial commutation rules from a computational point of view ([2,14,19,24,31] and references therein). The algebras considered are often realized as operator algebras and model linear functional equations. In that context, Gröbner bases of ideals (or modules), i.e. finite sets that both generate and provide a membership test to the ideal (or module), can be defined and computed by variants of Buchberger algorithm. We are dealing here with nonlinear differential equations. In this context the best representations that can be achieved in finite terms are valid outside some hypersurfaces [6,8,10,15,21,27,28,34–36,38]. We could speak of pseudo-Gröbner bases but we stick to the terminology of characteristic sets introduced by Ritt [36].

In Section 2, we outline on an example how the differential algebras we want to study arise. In Section 3, we define the ring of differential polynomials when the derivations neither commute nor satisfy commutation rules. In Section 4, we study the quotient of that formal ring with the relationships induced by the commutation rules of the derivations. We establish sufficient conditions for the quotient to be (algebraically) isomorphic to $\mathscr{F}[y_\alpha \mid y \in \mathscr{Y}, \alpha \in \mathbb{N}^m]$. As can be expected, the conditions bear on the coefficients c_{ijl} of the commutation rules. They are quite natural when written in terms of the appropriately defined bracket. In Section 5, we transport on $\mathscr{F}[y_\alpha \mid y \in \mathscr{Y}, \alpha \in \mathbb{N}^m]$ the appropriate differential structure and give explicit recursive definition of derivations. We can then define properly the differential polynomial ring with nontrivial commutation rules. In Section 6, we outline the constructions leading to a characteristic decomposition algorithm. Indeed, once a couple of fundamental properties are exhibited there is a close parallel between the constructions in classical differential algebra and the extension of it presented in this paper. Recent developments were given in great details in the tutorial [17] and we shall avoid repetition.

2. A motivating example

Consider the following differential system in two independant variables x and y and three dependant variables ϕ, ψ and s. It describes waves that would be othogonal. The underlying question, posed by Métivier, is to find the conditions on s for the system to have solutions:

$$S \begin{cases} s(\phi_{xx} + \phi_{yy}) + s_x\phi_x + s_y\phi_y + \phi = 0, \\ s(\psi_{xx} + \psi_{yy}) + s_x\psi_x + s_y\psi_y + \psi = 0, \\ \psi_x\phi_x + \psi_y\phi_y = 0. \end{cases}$$

To look at this system algebraically, we place ourselves in a *differential polynomial ring* with *differential indeterminates* $\{s, \phi, \psi\}$ and *derivations* $\{\partial/\partial x, \partial/\partial y\}$ [21,36]. The coefficient field $\mathbb{K}$ can be taken as $\mathbb{Q}$ or $\mathbb{C}$. To answer the question one could compute a *characteristic decomposition* of the radical differential ideal generated by the underlying differential polynomials w.r.t. a *ranking* that *eliminates* ψ and ϕ [17].

We can see that the system is rather symmetric and we want to exploit that fact to obtain the answer. One can check that the system is indeed left invariant by the following seven-dimensional Lie group action of the zeroth-order jet space. This symmetry was in fact computed with the help of the MAPLE *Desolv* package [9]. A group element g is determined by a 7-tuple of parameters $(t, \rho, a, b, \mu, \nu)$. Its pull back action on the coordinate functions $((x, y), (s, \phi, \psi))$ of $J^0(\mathbb{K}^2, \mathbb{K}^4)$ is given by the following expressions:

$$g^*x = \frac{(1-t^2)}{\rho(1+t^2)}x - \frac{2t}{\rho(1+t^2)}y + \frac{a}{\rho},$$

$$g^*y = \frac{2t}{\rho(1+t^2)}x + \frac{(1-t^2)}{\rho(1+t^2)}y + \frac{b}{\rho},$$

$$g^*s = \frac{s}{\rho^2}, \quad g^*\phi = \frac{\phi}{\mu}, \quad g^*\psi = \frac{\psi}{\nu}.$$

The system can thus be rewritten in terms of a set of *fundamental differential invariants* and the related two *invariant derivations* constructed by the moving frame method of [13].

A fundamental set of differential invariants $\{s_1, s_2, s_3, \psi_1, \psi_2, \phi_1, \phi_2\}$ was computed with the *Vessiot* package [1] and the *Groebner* library of MAPLE. The detailed computations are presented in [18]:

$$s_1^2 := \frac{s_x^2 + s_y^2}{4s},$$

$$s_2 := \frac{s_{xy}(s_y^2 - s_x^2) + s_x s_y (s_{xx} - s_{yy})}{8 s s_1^3}, \quad s_3 := \frac{s_x^2 s_{yy} + s_y^2 s_{xx} - 2 s_x s_y s_{xy}}{8 s s_1^3},$$

$$\psi_1 := \frac{s_y \psi_x - s_x \psi_y}{2 s_1 \psi}, \quad \psi_2 := \frac{s_x \psi_x + s_y \psi_y}{2 s_1 \psi},$$

$$\phi_1 := \frac{s_y \phi_x - s_x \phi_y}{2 s_1 \phi}, \quad \phi_2 := \frac{s_x \phi_x + s_y \phi_y}{2 s_1 \phi}.$$

We can write down the two invariant derivations in terms of $\partial/\partial x$ and $\partial/\partial y$

$$\begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} = \frac{\pm\sqrt{s(s_y^2 + s_x^2)}}{s_x^2 - s_y^2} \begin{pmatrix} -s_y & s_x \\ s_x & -s_y \end{pmatrix} \begin{pmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \end{pmatrix}.$$

The problem of computing the explicit expression for generating differential invariants $s_1, s_2, s_3, \phi_1, \phi_2, \psi_1, \psi_2$, on the one hand, and the invariant derivations δ_1, δ_2, on the other hand, actually separate problem. We can rewrite the system $\mathscr{S}$ in terms of those, compute the commutator of δ_1 and δ_2 and compute the differential relationships (syzygies) among the generating differential invariants with only the knowledge of the infinitesimal generators and a choice of a cross section to the orbits [13,20]. For that we extended the IVB program by Kogan [20]. We obtain on this example the commutator directly in terms of s_1, s_2, s_3

$$\delta_1\delta_2 - \delta_2\delta_1 = (s_3 - s_1)\delta_1 + s_2\delta_2.$$

The system can then be rewritten as

$$\mathscr{S}\begin{cases}\delta_1(\phi_1) + \delta_2(\phi_2) + \phi_1^2 + \phi_2^2 - \phi_1 s_2 + (s_1 + s_3)\phi_2 + 1 = 0,\\ \delta_1(\psi_1) + \delta_2(\psi_2) + \psi_1^2 + \psi_2^2 - \psi_1 s_2 + (s_1 + s_3)\psi_2 + 1 = 0,\\ \phi_1\psi_1 + \phi_2\psi_2 = 0\end{cases}$$

while the syzygies are

$$\mathscr{Z}\begin{cases}\delta_1(s_1) = s_1 s_2,\\ \delta_1(s_2) - \delta_2(s_3) = s_3^2 + s_2^2 - s_1 s_3,\\ \delta_1(\phi_2) - \delta_2(\phi_1) = \phi_1(s_3 - s_1) + \phi_2 s_2,\\ \delta_1(\psi_2) - \delta_2(\psi_1) = \psi_1(s_3 - s_1) + \psi_2 s_2.\end{cases}$$

What is suggested by this example is to consider the problem in the differential algebra where the set of differential indeterminates is $\mathscr{Y} = \{s_1, s_2, s_3, \phi_1, \phi_2, \psi_1, \psi_2\}$ and the set of derivations is $\Delta = \{\delta_1, \delta_2\}$. We consider the system $\mathscr{S} \cup \mathscr{Z}$. We face two obvious difficulties:

- the derivations do not commute, contrary to the basic assumption made in the classical differential algebra setting [17,21,36],
- the coefficients in the commutators are differential polynomials.

Let us note now two *lucky* properties[2] that are important to the particular problem presented above. First, s_1, s_2, s_3 depend only on the original function s and its derivatives. Thus a differential relationship in s_1, s_2, s_3 will mean a differential relationship on the original differential indeterminate s. Second, the coefficients of the commutator of δ_1 and δ_2 depend only on (s_1, s_2, s_3). So, if we use a block elimination ranking $s_1, s_2, s_3 \ll \phi_1, \phi_2, \psi_1, \psi_2$ we will be able to find additional relationships in (s_1, s_2, s_3) if there are any.

In this paper, we present the algebraic foundations for the treatment of the problem in terms of those new differential indeterminates and the non commuting new derivations.

3. Derivatives

In this section, we introduce a differential polynomial ring for strictly noncommuting derivations. In classical differential polynomial rings the indeterminates are indexed by *terms*. The indeterminates here are indexed by *words*. We shall be concerned, in next section, with the quotient of this differential polynomial ring by the relationships induced by the commutation rules on the derivations.

---

[2] We in fact made the appropriate choices in the moving frame construction so that those arise.

3.1. Words and terms

Let $m \in \mathbb{N}\setminus\{0\}$ and $\mathbb{N}_m = \{1, \dots, m\}$. We consider $\mathscr{W}_m$ the semi-group of *words* formed on $\mathbb{N}_m$: an element $I \in \mathscr{W}_m$ can be represented by an empty tuple () or a tuple $(i_1, i_2, \dots, i_p)$ for some $p \in \mathbb{N}\setminus\{0\}$ and $i_k \in \mathbb{N}_m$. The length of a word $I = (i_1, \dots, i_p)$ is $|I| = p$. For $I = (i_1, \dots, i_p)$ and $J = (j_1, \dots, j_q)$ two words of $\mathscr{W}_m$ we denote $I.J = (i_1, \dots, i_p, j_1, \dots, j_q)$ the *concatenation* of I and J. By extension, if $i_0 \in \mathbb{N}_m$ and $I = (i_1, i_2, \dots, i_p) \in \mathscr{W}_m$ we write $i_0.I$ for the element $(i_0, i_1, \dots, i_p)$ of $\mathscr{W}_m$. The word () is the neutral element of the concatenation and therefore $(\mathscr{W}_m, .)$ is a monoid.

We have a natural semi-group morphism λ from $\mathscr{W}_m$ to the commutative semi-group $\mathbb{N}^m$ that associates to the word $I = (i_1, \dots, i_p)$ the m-tuple $\alpha = (\alpha_1, \dots, \alpha_m)$ where α_k is the cardinal of the set $\{j \in \mathbb{N}_m \,|\, i_j = k\}$. $\mathbb{N}^m$ is isomorphic to the semi-group of terms $\mathscr{T}_m(\xi) = \{\xi_1^{\alpha_1} \dots \xi_m^{\alpha_m} \,|\, (\alpha_1, \dots, \alpha_m) \in \mathbb{N}^m\}$ in a set of m indeterminates $\xi = \{\xi_1, \dots, \xi_m\}$. To avoid confusion with the different types of tuples, we call an element of $\mathbb{N}^m$ a *term* and we shall note $\mathscr{T}_m = \mathbb{N}^m$ for a unified notation system. As a general rule, words will be denoted by capital letter while terms will be denoted by a Greek lower case letter.

A word $(i_1, \dots, i_p)$ is *monotone* if $i_1 \leqslant i_2 \leqslant \dots \leqslant i_p$. We note $\mathscr{M}_m$ the set of monotone words. The restriction of λ to $\mathscr{M}_m$ is one-to-one. For a word $I \in \mathscr{W}_m$ we shall note $\overline{I}$ the monotone word that has the same image as I by λ.

Example 3.1. If $m = 4$ and $I = (1, 4, 2, 1, 4, 1) \in \mathscr{W}_m$ then $\lambda(I) = (3, 1, 0, 2) \in \mathscr{T}_m$ and $\overline{I} = (1, 1, 1, 2, 4, 4) \in \mathscr{M}_m \subset \mathscr{W}_m$.

We recall the definition of the *lexicographical order* on $\mathscr{W}_m$. Let $I = (i_1, \dots, i_p)$ and $J = (j_1, \dots, j_q)$ be elements of $\mathscr{W}_m$

$$I \prec_{\text{lex}} J \text{ iff } \begin{cases} \exists\, K \in \mathscr{W}_m, \ \ J = I.K \text{ and } |K| > 0 \\ \text{or} \\ \exists\, K, I_1, J_1 \in \mathscr{W}_m, \ \ i, j \in \mathbb{N}_m \text{ s.t. } I = K.i.I_1 \\ \quad \text{and } J = K.j.J_1 \text{ and } i < j. \end{cases}$$

A monotone word is lower, with respect to the lexicographical order, to any word obtained from it by permutation of the components.

3.2. Indexed indeterminates and ranking

To a finite set $\mathscr{Y}$ of indeterminates we associate two infinite sets of indexed indeterminates, on the one hand, the ones indexed by terms $\mathscr{T}_m(\mathscr{Y}) = \{y_\alpha \,|\, y \in \mathscr{Y}, \ \alpha \in \mathscr{T}_m\}$ are called *derivatives* and, on the other hand, the ones indexed by words $\mathscr{W}_m(\mathscr{Y}) = \{y_I \,|\, y \in \mathscr{Y}, \ I \in \mathscr{W}_m\}$ are called *word derivatives*. Let $\mathbb{K}$ be a field (of characteristic zero). We shall consider the polynomial rings $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ and $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$.

We shall often abbreviate elements of $\mathscr{W}_m(\mathscr{Y})$. For instance $y_{()}$ becomes y and $y_{(i,j,k)}$ becomes $y_{i.j.k}$.

The semi-group morphism λ is extended to a $\mathbb{K}$-algebra morphism

$$\lambda : \mathbb{K}[\mathscr{W}_m(\mathscr{Y})] \to \mathbb{K}[\mathscr{T}_m(\mathscr{Y})],$$

$$y_I \mapsto y_{\lambda(I)}.$$

For clarity we shall also use the subset $\mathscr{M}_m(\mathscr{Y}) = \{y_I \mid y \in \mathscr{Y},\ I \in \mathscr{M}_m\} \subset \mathscr{W}_m(\mathscr{Y})$ of indeterminates indexed by monotone words that we may call *monotone derivatives*. The associated polynomial ring $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ through λ.

A *ranking*[3] on $\mathscr{T}_m(\mathscr{Y})$ is a total order $\prec$ on $\mathscr{T}_m(\mathscr{Y})$ such that

- $y_\alpha \prec y_{\alpha+\beta},\ \forall \alpha, \beta \in \mathscr{T}_m,\ \forall y \in \mathscr{Y}.$
- $y_\alpha \prec z_\beta \Rightarrow y_{\alpha+\gamma} \prec z_{\beta+\gamma},\ \forall \alpha, \beta, \gamma \in \mathscr{T}_m,\ \forall y, z \in \mathscr{Y}.$

A ranking on $\mathscr{T}_m(\mathscr{Y})$ is extended to $\mathscr{W}_m(\mathscr{Y})$ by lexicographical order, i.e.

$$y_I \prec z_J \Leftrightarrow \begin{cases} \lambda(y_I) \prec_{\mathscr{T}_m} \lambda(z_J) \\ \text{or} \\ \lambda(y_I) = \lambda(z_J) \text{ and } I \prec_{\text{lex}} J. \end{cases}$$

We shall speak directly of a *ranking on* $\mathscr{W}_m(\mathscr{Y})$. This assumes that if $y_{\overline{I}} \prec z_{\overline{J}}$ then $y_I \prec z_J$.

Proposition 3.2. *A ranking on $\mathscr{W}_m(\mathscr{Y})$ is a well order, i.e. any strictly decreasing sequence of elements of $\mathscr{W}_m(\mathscr{Y})$ is finite.*

Proof. A ranking on $\mathscr{T}_m(\mathscr{Y})$ refines the product order on $\mathbb{N}^m$. By Dickson's lemma it is a well order. As the preimage of an element of $\mathscr{T}_m(\mathscr{Y})$ by λ has a finite cardinal, the ranking on $\mathscr{W}_m(\mathscr{Y})$ is also a well order. $\square$

Given a ranking on $\mathscr{W}_m(\mathscr{Y})$ we can define, as is usual in differential algebra and in the theory of triangular sets [16,17], for any element $p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\setminus\mathbb{K}$

- $\mathsf{lead}(p)$: the *leader* of p is the highest ranking (word) derivative in p,
- $\mathsf{rank}(p)$: the *rank* of p is the leader raised at the highest power appearing in p,
- $\mathsf{init}(p)$: the *initial* of p is the coefficient of $\mathsf{rank}(p)$ in p considered as a polynomial in $\mathsf{lead}(p)$,
- $\mathsf{sep}(p)$: the *separant* of p is the formal derivative of p w.r.t. $\mathsf{lead}(p)$,
- $\mathsf{tail}(p)$: is $p - \mathsf{init}(p)\,\mathsf{rank}(p)$.

A *triangular set* is a set of elements of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\setminus\mathbb{K}$ the leaders of which are pairwise distinct.

A ranking on $\mathscr{W}_m(\mathscr{Y})$ induces a pre-order on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$

$$p \prec q \text{ if } \begin{cases} p \in \mathbb{K} \text{ and } q \notin \mathbb{K}, \\ \text{or } \mathsf{lead}(p) \prec \mathsf{lead}(q), \\ \text{or } \mathsf{lead}(p) = \mathsf{lead}(q) \text{ and } \deg(p, \mathsf{lead}(p)) < \deg(q, \mathsf{lead}(p)). \end{cases}$$

[3] It amounts to an admissible term ordering on the free module $\mathbb{K}[\xi]^{|\mathscr{Y}|}$.

In classical differential algebra, orderly ranking are especially important in connection with formal integrability and the differential dimension polynomial or the Cartan characters. We introduce here semi-orderly rankings as they appear necessary.

Definition 3.3. A ranking is *orderly*[4] if whenever $|I| < |J|$, for $I, J \in \mathscr{W}_m$, then $y_I \prec z_J$ for any $y, z \in \mathscr{Y}$. A ranking is *semi-orderly* if whenever $|I| < |J|$ then $y_I \prec y_J$, for all $y \in \mathscr{Y}$.

3.3. Derivations

A derivation δ on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is a map from $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ to $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ that is $\mathbb{K}$-linear and s.t. $\delta(ab) = a\delta(b) + \delta(a)b$. On $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ we define m derivations $\delta_1, \ldots, \delta_m$ by

$$\delta_i(y_I) = y_{i.I}, \quad \forall i \in \mathbb{N}_m, \forall y_I \in \mathscr{W}_m(\mathscr{Y}).$$

If $\prec$ is a ranking on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ then we retrieve the classical conditions of compatibility with derivation:

- $y_I \prec \delta_i(y_I)$ for all $i \in \mathbb{N}_m$ and $y_I \in \mathscr{W}_m(\mathscr{Y})$,
- $y_I \prec z_J \Rightarrow \delta_i(y_I) \prec \delta_i(z_J)$ for all $i \in \mathbb{N}_m$ and $y_I, z_J \in \mathscr{W}_m(\mathscr{Y})$.

For $I = (i_1, \ldots, i_p) \in \mathscr{W}_m$ we shall denote $\delta_{i_1} \circ \cdots \circ \delta_{i_p}$ by δ^I. For $\alpha \in \mathscr{T}_m$, δ^α denotes δ^I where I is the only monotone word with $\lambda(I) = \alpha$. Obviously $\delta^{\alpha+\beta} \neq \delta^\alpha \circ \delta^\beta$.

4. Normalization of derivatives

When derivations commute all derivation operators can be expressed in terms of $\delta_1^{\alpha_1} \ldots \delta_m^{\alpha_m}$ and therefore the derivatives need only be indexed by terms. If the coefficients of the commutation rules are constants, we can achieve a similar basis for the set of derivation operators thanks to the Poincaré–Birkhoff–de Witt theorem. In our case the derivatives occur in the commutation rules. Working at the operator level is thus not an option. We need to find a basis for the derivatives.

In this section, we show that under some natural conditions $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Omega)$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ where Ω is the set of the *commutation relationships* implied by the *commutation rules* on the derivations. The result is reminiscent of the PBW theorem but cannot be obtained from it. In the same line, Mora [32] was a source of understanding and inspiration but we could not apply the results directly.

The proof is as follows. We first select a subset Γ of Ω. Provided there exists a ranking that is compatible with the commutation rules, the quotient of the formal differential polynomial ring by (Γ) is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$. Also Γ defines a constructive normalization. We then prove that $(\Omega) = (\Gamma)$ under some natural conditions on the commutation rules.

[4] The definition of [21] is here adapted to word derivatives.

4.1. Commutation rules and relationships

Let $\mathscr{M}$ be the free $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$-module $\oplus_{i=1}^m \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\bar{\delta}_i$. To a family $\mathscr{C}=\{c_{ijk} \mid i, j, k \in \mathbb{N}_m\}$ of elements of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ we associate the $\mathbb{K}$-bilinear map $[\cdot,\cdot] : \mathscr{M} \to \mathscr{M}$ with the following rules for $i, j \in \mathbb{N}_m$ and $p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$:

$$[\bar{\delta}_i, \bar{\delta}_j] = \sum_{k=1}^m c_{ijk}\bar{\delta}_k,$$

$$[p\bar{\delta}_i, \bar{\delta}_j] = p[\bar{\delta}_i, \bar{\delta}_j] - \delta_j(p)\bar{\delta}_i,$$

$$[\bar{\delta}_i, p\bar{\delta}_j] = p[\bar{\delta}_i, \bar{\delta}_j] + \delta_i(p)\bar{\delta}_j.$$

We can consider the associative $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$-algebra generated by $\bar{\Delta}=\{\bar{\delta}_1, \ldots, \bar{\delta}_2\}$, written $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\langle\bar{\Delta}\rangle$, in which the inner product (composition) $\circ$ satisfies

$$\bar{\delta}_i \circ (p \cdot \theta) = \delta_i(p) \cdot \theta + p \cdot \bar{\delta}_i \circ \theta, \quad \forall p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})], \forall \theta \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\langle\bar{\Delta}\rangle,$$

where $\cdot$ is the product with an element of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$.

We note $\{\cdot,\cdot\}$ the commutator in $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\langle\bar{\Delta}\rangle$ that is

$$\{\psi, \phi\} = \psi \circ \phi - \phi \circ \psi, \quad \psi, \phi \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\langle\bar{\Delta}\rangle.$$

Let $\gamma = \{\gamma_{ij}\}_{i,j\in\mathbb{N}_m}$ be the elements in $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\langle\bar{\Delta}\rangle$ defined as

$$\gamma_{ij} = \{\bar{\delta}_i, \bar{\delta}_j\} - [\bar{\delta}_i, \bar{\delta}_j].$$

Mapping $\bar{\delta}_i$ onto δ_i, the elements of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\langle\bar{\Delta}\rangle$ can be considered as operators on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$. The actions of the γ_{ij} for instance is given by

$$\gamma_{ij} : \mathbb{K}[\mathscr{W}_m(\mathscr{Y})] \to \mathbb{K}[\mathscr{W}_m(\mathscr{Y})],$$

$$p \mapsto \delta_i(\delta_j(p)) - \delta_j(\delta_i(p)) - \sum_{l=1}^m c_{ijl}\delta_l(p).$$

In the following, we shall simply write δ_i for $\bar{\delta}_i$ and the multiplications $\circ$ and $\cdot$ shall be omitted.

$\{\delta_i, \delta_j\}$, $[\delta_i, \delta_j]$ and therefore γ_{ij} are derivations on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$. We shall refer to the γ_{ij} as the *commutation rules* of the derivations $\{\delta_1, \ldots, \delta_m\}$.

A ranking on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is *compatible* with the commutation rules if for all $y_I \in \mathscr{W}_m(\mathscr{Y})$ and $i, j \in \mathbb{N}_m$ we have $[\delta_i, \delta_j](y_I) \prec y_{\overline{j.i.I}}$. Note that it is not enough to check that $[\delta_i, \delta_j](y) \prec y_{\overline{j.i}}$.

Example 4.1. Assume $c_{211}=y_{1.1.2}\,y_2=-c_{121}$ and $c_{212}=y_{1.1.2}\,y_1=-c_{122}$. Then $[\delta_2, \delta_1](y) = 0$ but $[\delta_2, \delta_1](y_2) = y_{1.1.2}\,(y_2 y_{1.2} - y_1 y_{2.2})$.

Nonetheless, any orderly ranking is compatible if the coefficients c_{ijl} involve no derivatives of order more than 1.

Let Ω be the set of all the generated *commutation relationships* induced by the γ_{ij}. That is

$$\Omega = \{\delta^J \gamma_{ij}(p) \mid i, j \in \mathbb{N}_m, \;\; J \in \mathscr{W}_m, \;\; p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\}.$$

We shall show that under some natural conditions on the coefficients $\{c_{ijl}\}_{i,j,l\in\mathbb{N}_m}$ of the commutation rules we have an isomorphism

$$\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Omega) \cong \mathbb{K}[\mathscr{T}_m(\mathscr{Y})].$$

4.2. Reduction of word derivatives to monotone derivatives

In this section, we exhibit a subset Γ of Ω such that $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Gamma)$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$. This set actually induces a normal form algorithm modulo (Γ).

Let

$$\Gamma = \{\delta^J \gamma_{ji}(y_I) \mid i, j \in \mathbb{N}_m, I, J \in \mathscr{W}_m \text{ s.t. } j > i, \;\; i.I \in \mathscr{M}_m\}.$$

Γ is stable under the actions of $\delta_1, \ldots, \delta_m$ and thus so is the generated ideal (Γ). This set is constructed so that we have the following property.

Lemma 4.2. *Assume $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is endowed with a ranking that is compatible with the commutation rules. Γ is a triangular set. Its set of leaders is the set of all the non monotone derivatives.*

Proof. For a compatible ranking the leader of $\delta^J \gamma_{ji}(y_I)$ is $y_{J.j.i.I}$ when $j > i$. If $K = (k_1, \ldots, k_p) \in \mathscr{W}_m$ is not monotone then y_K is the leader of a single element of Γ. Indeed taking

- t to be the smallest integer s.t. $(k_t, k_{t+1}, \ldots, k_p)$ is monotone. Note that $2 \leqslant t \leqslant p$.
- $j = k_{t-1}$ and $i = k_t$. Note that $j > i$ since otherwise $(k_{t-1}, k_t, \ldots, k_p)$ would be monotone.
- $I = ()$ if $t = p$ or $I = (k_{t+1}, \ldots, k_p)$ otherwise. Note that $i.I$ is monotone.
- $J = ()$ if $t = 2$ or $J = (k_1, \ldots, k_{t-2})$ otherwise

y_K is the leader of $\delta^J \gamma_{ji}(y_I)$. □

Lemma 4.3. *Assume $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is endowed with a ranking that is compatible with the commutation rules. Any element of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is equal modulo (Γ) to a polynomial of $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$.*

Proof. We proceed by contradiction. Consider the polynomials of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ the class of which modulo (Γ) does not include a polynomial in $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$. Take one of those in which the highest ranking non monotone derivative y_K is minimal. Call that polynomial p. This non monotone derivative is the leader of an element γ of Γ. Since γ is of degree one in y_K with initial 1, p can be rewritten modulo γ into a q s.t. the only non monotone derivatives that q contains rank lower than y_K. By the choice of p and y_K, q must be equal modulo (Γ) to a polynomial in monotone derivatives only. So must p then. □

The result is actually constructive.

Proposition 4.4. *There is an algorithm, normal_Γ, that computes for any $p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ a polynomial $\mathit{normal}_\Gamma(p) \in \mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ such that $\mathit{normal}_\Gamma(p) - p \in (\Gamma)$. We have the property that $\mathit{normal}_\Gamma(y_I) - y_{\overline{I}}$ ranks lower than $y_{\overline{I}}$.*

Proof. The algorithm consists, at each step, in rewriting the highest ranking non monotone derivative in the polynomial with the only element of Γ having this derivative as leader. As rankings are well orders (Proposition 3.2), the process terminates. At each step a y_I is actually replaced by a $y_J + q$ where $\lambda(I) = \lambda(J)$ and $J \prec_{\text{lex}} I$ and q involves only derivatives ranking lower than $\overline{I}$. The second statement is clear. □

We shall see in next lemma that the algorithm indeed provides a normal form modulo Γ: $\mathit{normal}_\Gamma(p)$ is the only element of $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ that belongs to the class of p modulo (Γ).

Lemma 4.5. *Assume $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is endowed with a ranking that is compatible with the commutation rules. (Γ) is a prime ideal and the set of monotone derivatives, $\mathscr{M}_m(\mathscr{Y})$, forms a maximally independent set modulo (Γ).*

Proof. Consider $\Gamma' = \{\mathsf{lead}(\gamma) - \mathit{normal}_\Gamma(\gamma - \mathsf{lead}(\gamma)) \mid \gamma \in \Gamma\}$. Evidently $(\Gamma) = (\Gamma')$. The only non monotone derivative in an element of Γ' is its leader and it appears with degree one and initial 1.

If A is a finite subset of Γ', A is a triangular set and actually a regular chain [16, Section 5]. The set of leaders of A is exactly the set of the non monotone derivatives appearing in A. There is a finite number of monotone derivatives appearing in A. By linearity of the elements of A in their leaders and the properties of regular chains, $(A) = (A) : I_A^\infty$ is prime and that finite number of monotone derivatives is a transcendence basis for it [16, Section 4].

If p belongs to (Γ') there is a finite subset A of Γ' s.t. $p \in (A)$. From what precedes, p cannot involve monotone derivatives only. Assume $p_1 p_2 \in (\Gamma')$. Then there exists a finite subset A of Γ' s.t. $p_1 p_2 \in (A)$. As (A) is prime either $p_1 \in (A) \subset (\Gamma')$ or $p_2 \in (A) \subset (\Gamma')$. Consequently $(\Gamma') = (\Gamma)$ is prime and the set of monotone derivatives is algebraically independent modulo (Γ). As any other word derivative is algebraically dependent over those modulo (Γ') the conclusion follows. □

Proposition 4.6. *Assume $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ is endowed with a ranking that is compatible with the commutation rules. The algorithm normal_Γ expresses a surjective morphism from $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ to $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ the kernel of which is (Γ). $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Gamma)$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$.*

Proof. By the previous two lemmas, for any $p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$, there is a unique element in the class of p modulo (Γ) that belongs to $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$. That element is $\mathit{normal}_\Gamma(p)$. It follows that normal_Γ is a ring epimorphism from $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ to $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$. Its kernel is (Γ).

Recall that $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ through λ. Thus $\lambda \circ \mathit{normal}_\Gamma : \mathbb{K}[\mathscr{W}_m(\mathscr{Y})] \to \mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ is a ring epimorphism with (Γ) as kernel. It follows that $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Gamma)$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$. □

For visualization we have the following commuting diagram:

$$\begin{array}{ccc} \mathbb{K}[\mathscr{W}_m(\mathscr{Y})] & \xrightarrow{\mathsf{normal}_\Gamma} & \mathbb{K}[\mathscr{M}_m(\mathscr{Y})] \\ \downarrow & & \downarrow \lambda \\ \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Gamma) & \xrightarrow{\sim} & \mathbb{K}[\mathscr{T}_m(\mathscr{Y})] \end{array}$$

4.3. Quotient ring

We shall prove in this section that given some natural conditions on the commutation rules the ideal $(\Gamma)=(\Omega)$, i.e. that (Γ) contains all the commutation relationships on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$.

Proposition 4.7. *Assume that $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ can be endowed with a semi-orderly ranking $\prec$ that is compatible with the commutation rules. If the commutation rules satisfy*

1. $[\delta_i, \delta_j] = -[\delta_j, \delta_i]$ *for all* $i, j \in \mathbb{N}_m$,
2. $[[\delta_i, \delta_j], \delta_k] + [[\delta_j, \delta_k], \delta_i] + [[\delta_k, \delta_i], \delta_j] = 0$ *for all* $i, j, k \in \mathbb{N}_m$,

where the coefficients are understood modulo (Γ), then $\delta^I \gamma_{ij}(p) \in (\Gamma)$, for all $p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$, $i, j \in \mathbb{N}_m$ and $I \in \mathscr{W}_m$.

Proof. As (Γ) is stable under the actions of the derivations $\Delta=\{\delta_1, \ldots, \delta_m\}$, it is enough to prove that $\gamma_{ij}(p) \in (\Gamma)$. Furthermore, as the γ_{ij} are derivations on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$, it is enough to prove that $\gamma_{ij}(y_I) \in (\Gamma)$ for all $i, j \in \mathbb{N}_m$ and all $y_I \in \mathscr{W}_m(\mathscr{Y})$.

As $[.,.]$ is assumed to be anti-symmetric we need only to consider the case $i > j$. The proof is by induction.

The result is true for any $y \in \mathscr{Y}$ and $I = ()$ as $\gamma_{ij}(y) \in \Gamma$ by construction.

Take $(i, j) \in \mathbb{N}_m \times \mathbb{N}_m$ such that $i > j$, $y \in \mathscr{Y}$ and $I \in \mathscr{W}_m$ such that $I \neq ()$. Assume that for any $(a, b, J, z) \in \mathbb{N}_m \times \mathbb{N}_m \times \mathscr{W}_m \times \mathscr{Y}$ such that $z_{a.b.J} \prec y_{i.j.I}$ we have $\gamma_{ab}(z_J) \in (\Gamma)$. We shall show that $\gamma_{ij}(y_I) \in (\Gamma)$.

As the ranking is compatible with the commutation rules and $[.,.]$ is anti-symmetric, we can write

$$y_I \equiv y_{\overline{I}} + \text{terms ranking lower than} y_{\overline{I}} \mod (\Gamma).$$

By induction hypothesis and because (Γ) is stable by derivation we have

$$\gamma_{ij}(y_I) \equiv \gamma_{ij}(y_{\overline{I}}) \bmod (\Gamma).$$

We have to consider two cases. Either $j \preceq_{\text{lex}} \overline{I}$ or the contrary. In the first case $\gamma_{ij}(y_{\overline{I}}) \in \Gamma$ by construction. In the second case, there is $k \in \mathbb{N}_m$ and $K \in \mathscr{W}_m$ s.t. $\overline{I} = k.K$ with $j > k$. We have

$$\begin{aligned} \{\delta_k, \gamma_{ij}\} + \{\delta_i, \gamma_{jk}\} + \{\delta_j, \gamma_{ki}\} &= [[\delta_i, \delta_j], \delta_k] + [[\delta_j, \delta_k], \delta_i] + [[\delta_k, \delta_i], \delta_j] \\ &\quad + \sum_{l=1}^{m} c_{ijl}\gamma_{lk} + c_{jkl}\gamma_{li} + c_{kil}\gamma_{lj}. \end{aligned}$$

As

$$\{\delta_k, \gamma_{ij}\} = \{\delta_k, \{\delta_i, \delta_j\}\} + \{[\delta_i, \delta_j], \delta_k\}$$

and

$$\{[\delta_i, \delta_j], \delta_k\} = \sum_{l=1}^{m} c_{ijl}\gamma_{lk} + [[\delta_i, \delta_j], \delta_k]$$

while the commutator $\{\cdot,\cdot\}$ satisfies a Jacobi identity. Using the Jacobi identity on $[\cdot,\cdot]$ of the hypothesis we have thus

$$\delta_k\gamma_{ij} - \gamma_{ij}\delta_k + \delta_i\gamma_{jk} - \gamma_{jk}\delta_i + \delta_j\gamma_{ki} - \gamma_{ki}\delta_j = \sum_{l=1}^{m} c_{ijl}\gamma_{lk} + c_{jkl}\gamma_{li} + c_{kil}\gamma_{lj}.$$

We can write

$$\begin{aligned}\gamma_{ij}(y_{\bar{I}}) &= \gamma_{ij}(y_{k.K}) = \gamma_{ij}\delta_k(y_K)\\ &= \delta_k\gamma_{ij}(y_K) + \delta_i\gamma_{jk}(y_K) + \delta_j\gamma_{ki}(y_K) - \gamma_{jk}\delta_i(y_K) - \gamma_{ki}\delta_j(y_K)\\ &\quad - \sum_{l=1}^{m} c_{ijl}\gamma_{lk}(y_K) + c_{jkl}\gamma_{li}(y_K) + c_{kil}\gamma_{lj}(y_K)\\ &= \delta_k\gamma_{ij}(y_K) + \delta_i\gamma_{jk}(y_K) + \delta_j\gamma_{ki}(y_K) - \gamma_{jk}(y_{i.K}) - \gamma_{ki}(y_{j.K})\\ &\quad - \sum_{l=1}^{m} c_{ijl}\gamma_{lk}(y_K) + c_{jkl}\gamma_{li}(y_K) + c_{kil}\gamma_{lj}(y_K).\end{aligned}$$

As $i > j > k$, $j.k.i.K \prec_{\text{lex}} i.j.k.K$ and $k.i.j.K \prec_{\text{lex}} i.j.k.K$. Thus $y_{j.k.i.K} \prec y_{i.j.\bar{I}} \preceq y_{i.j.I}$ and $y_{k.i.j.K} \prec y_{i.j.I}$. By induction hypothesis $\gamma_{jk}(y_{i.K})$ and $\gamma_{ki}(y_{j.K})$ belong to (Γ).

As $|i.j.K|$, $|j.k.K|$, $|k.i.K|$, on the one hand, and $|l.k.K|$, $|l.i.K|$, $|l.j.K|$, on the other hand, are smaller than $|i.j.k.K| = |i.j.I|$, the associated derivatives $y_{i.j.K}$, $y_{j.k.K}$, $y_{k.i.K}$, $y_{l.k.K}$, $y_{l.i.K}$, $y_{l.j.K}$ rank lower than $y_{i.j.I}$ since the ranking is assumed to be semi-orderly. By induction hypothesis we can conclude that $\gamma_{ij}(y_K)$, $\gamma_{jk}(y_K)$, $\gamma_{ki}(y_K)$, $\gamma_{lk}(y_K)$, $\gamma_{li}(y_K)$, $\gamma_{lj}(y_K)$ belong to (Γ). As (Γ) is stable under the action of the elements of Δ, $\delta_k\gamma_{ij}(y_K)$, $\delta_i\gamma_{jk}(y_K)$ and $\delta_j\gamma_{ki}(y_K)$ belong to (Γ) and the conclusion follows. $\square$

The two conditions expressed in terms of the bracket are actually conditions on the coefficients $\{c_{ijl}\}_{i,j,l\in\mathbb{N}_m}$. The first is $c_{ijl} = -c_{jil} \bmod (\Gamma)$ for all $i, j, l \in \mathbb{N}_m$. The second condition is

$$\begin{aligned}&\sum_{\mu=1}^{m} c_{ij\mu}c_{\mu kl} + c_{jk\mu}c_{\mu il} + c_{ki\mu}c_{\mu jl}\\ &\quad = \delta_k(c_{ijl}) + \delta_i(c_{jkl}) + \delta_j(c_{kil}) \mod (\Gamma), \quad \forall i, j, k, l \in \mathbb{N}_m.\end{aligned}$$

If any of those conditions are not satisfied, there is a polynomial of $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ that belongs to (Ω). For instance, if the conditions $c_{ijl} + c_{jil} = 0$ are not all satisfied then $\gamma_{ij}(y) + \gamma_{ji}(y) = \sum_{l=1}^{m}(c_{ijl} - c_{jil})y_l \in \Omega$. Assuming that the c_{ijl} are constants, it implies that there is a

$\mathbb{K}$-linear dependency of the first-order derivatives implied by the commutation relationship. Similarly, if the second condition is not satisfied, we shall find that some polynomial of $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ belong to (Ω). For example, if all the c_{ijk} belong to $\mathbb{K}$ and do not satisfy the condition $\sum_{\mu=1}^{m} c_{ij\mu}c_{\mu kl} + c_{jk\mu}c_{\mu il} + c_{ki\mu}c_{\mu jl} = 0 \mod (\Gamma)$ then there is a non trivial linear combination of the y_i that belongs to (Ω). We thus understand that those conditions are highly desirable.

Note that a sufficient condition for the existence of a semi-orderly ranking that is compatible with the commutation rules is that all the c_{ijl} are of order one or less: $\{c_{ijl}\}_{i,j,l} \subset \mathbb{K}[y_I \mid y \in \mathscr{Y}, |I| \leqslant 1]$.

Taking $\hat{\lambda} = \lambda \circ \mathit{normal}_\Gamma$, we may summarize the results of this section in the following theorem.

Theorem 4.8. *If $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]$ can be endowed with a semi-orderly ranking $\prec$ that is compatible with the commutation rules and if the commutation rules satisfy*

1. $[\delta_i, \delta_j] = -[\delta_j, \delta_i]$ *for all* $i, j \in \mathbb{N}_m$,
2. $[[\delta_i, \delta_j], \delta_k] + [[\delta_j, \delta_k], \delta_i] + [[\delta_k, \delta_i], \delta_j] = 0$ *for all* $i, j, k \in \mathbb{N}_m$,

then $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Omega)$ is isomorphic to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$, where Ω is the set of commutation relationships induced by the commutation rules on the derivations

$$\Omega = \{\delta^J(\{\delta_i, \delta_j\} - [\delta_i, \delta_j])(p) \mid i, j \in \mathbb{N}_m,\ J \in \mathscr{W}_m,\ p \in \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]\}.$$

Furthermore, there is an algorithm to implement the epimorphism $\hat{\lambda} : \mathbb{K}[\mathscr{W}_m(\mathscr{Y})] \to \mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ s.t.

- *the kernel of $\hat{\lambda}$ is (Ω),*
- *the restriction of $\hat{\lambda}$ to $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ is λ,*
- *$\hat{\lambda}(y_I) - \lambda(y_I)$ ranks less than $\lambda(y_I)$,*
- *the following diagram commutes*:

$$\begin{tikzcd}
\mathbb{K}[\mathscr{W}_m(\mathscr{Y})] \arrow[rr, "\hat{\lambda}"] \arrow[dr] & & \mathbb{K}[\mathscr{T}_m(\mathscr{Y})] \\
& \mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Omega) \arrow[ur, "\sim"] &
\end{tikzcd}$$

5. Differential polynomial rings with nontrivial commutation rules

In this section we give the definition for a differential polynomial ring with non commuting derivations. The isomorphism exhibited in the previous section allows to define the sought differential structure. We first make explicit the action of the derivations. We will then be able to shift back to the usual notations of classical differential algebra.

The fact that we need to define (recursively) the actions of the derivations on the derivatives and the existence of an admissible ranking account for major differences for defining a differential polynomial ring with non commuting derivations compared to classical differential algebra.

5.1. Derivations and derivation operators

We take Ω, λ and $\hat{\lambda}$ as in previous section. As (Ω) is stable under the actions of $\Delta = \{\delta_1, \ldots, \delta_m\}$, Δ also defines derivations on $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Omega)$. We can then transport the differential structure of $\mathbb{K}[\mathscr{W}_m(\mathscr{Y})]/(\Omega)$ to $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ through the isomorphism. On $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ the derivations $\tilde{\Delta} = \{\tilde{\delta}_1, \ldots, \tilde{\delta}_m\}$ are defined by

$$\tilde{\delta}_i(p) = \hat{\lambda}(\delta_i(\mu(p))), \quad \forall p \in \mathbb{K}[\mathscr{T}_m(\mathscr{Y})],$$

where $\mu : \mathbb{K}[\mathscr{T}_m(\mathscr{Y})] \to \mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$ is the inverse of the restriction of λ, or equivalently $\hat{\lambda}$, to $\mathbb{K}[\mathscr{M}_m(\mathscr{Y})]$. We have $\hat{\lambda} \circ \delta_i = \tilde{\delta}_i \circ \hat{\lambda}$ and $\tilde{\delta}_i \tilde{\delta}_j(p) = \tilde{\delta}_j \tilde{\delta}_i(p) + \sum_{i=1}^{m} \hat{\lambda}(c_{ijl}) \tilde{\delta}_l(p)$, for all $p \in \mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$.

We can nonetheless compute the action of the derivations in a more direct way.

Let $\hat{c}_{ijl} = \hat{\lambda}(c_{ijl})$. On $\mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ consider the set of derivations $\hat{\Delta} = \{\hat{\delta}_1, \ldots, \hat{\delta}_m\}$ so that $\hat{\delta}_i|_{\mathbb{K}} = 0$ and defined on the $y_\alpha \in \mathscr{T}_m(\mathscr{Y})$ recursively as follows:

$$\hat{\delta}_i(y_\alpha) = \begin{cases} y_{\alpha+\varepsilon_i} & \text{if } \alpha_1 = \cdots = \alpha_{i-1} = 0, \\ \hat{\delta}_j \hat{\delta}_i(y_{\alpha-\varepsilon_j}) + \sum_{l=1}^{m} \hat{c}_{ijl} \hat{\delta}_l(y_{\alpha-\varepsilon_j}) & \text{where } j < i \text{ is s.t. } \alpha_j > 0 \\ & \text{while } \alpha_1 = \cdots = \alpha_{j-1} = 0, \end{cases}$$

where ε_i is the element of $\mathbb{N}^m$ having only 0 as components except at the ith position where there is a 1.

Proposition 5.1. $\hat{\lambda} \circ \delta_i = \hat{\delta}_i \circ \hat{\lambda}$, *for all* $i \in \mathbb{N}_m$.

Proof. We only need to prove that $\hat{\lambda}\hat{\delta}_i(y_I) = \hat{\delta}_i \hat{\lambda}(y_I)$ for all $y_I \in \mathscr{W}_m(\mathscr{Y})$. The proof is by induction. The base cases are immediate: for any $y \in \mathscr{Y}$ and $i \in \mathbb{N}_m$, $\hat{\lambda}(\hat{\delta}_i y) = \hat{\lambda}(y_i) = y_{\varepsilon_i} = \hat{\delta}_i \hat{\lambda}(y)$.

Assume now that $\hat{\lambda}(\hat{\delta}_k z_K) = \hat{\delta}_k \hat{\lambda}(z_K)$ for all $(k, K, z) \in \mathbb{N}_m \times \mathscr{W}_m \times \mathscr{Y}$ such that $z_{k.K} < y_{i.I}$ for some $(i, I, y) \in \mathbb{N}_m \times \mathscr{W}_m \times \mathscr{Y}$. We shall show the result is also true for (i, I, y).

Consider first the case where I is monotone. Take $\alpha = \lambda(I) \in \mathscr{T}_m$ and $j \in \mathbb{N}_m$, $J \in \mathscr{M}_m$ s.t. $I = j.J$. If $i \leqslant j$ then $i.I$ is monotone and $\alpha_1 = \cdots = \alpha_{i-1} = 0$ so that

$$\hat{\lambda}(\delta_i y_I) = \hat{\lambda}(y_{i.I}) = y_{\alpha+\varepsilon_i} = \hat{\delta}_i y_\alpha = \hat{\delta}_i \hat{\lambda}(y_I).$$

In the other case, when $i > j$, we have $\alpha_1 = \cdots = \alpha_{j-1} = 0$. We can write

$$\delta_i y_I = \delta_i \delta_j y_J \equiv \delta_j \delta_i y_J + \sum_{l=1}^{m} c_{ijl} \delta_l(y_J) \bmod (\Omega),$$

so that

$$\hat{\lambda}(\delta_i y_I) = \hat{\lambda}(\delta_j y_{i.J}) + \sum_{l=1}^{m} \hat{c}_{ijl}\hat{\lambda}(\delta_l y_J).$$

Since $j.i.J \prec_{\text{lex}} i.j.J = i.I$ we have $y_{j.i.J} \prec y_{i.I}$ so that by induction hypothesis $\hat{\lambda}(\delta_j y_{i.J}) = \delta_j \hat{\lambda}(\delta_i y_J)$. As the ranking is semi-orderly $y_{i.J}$ and $y_{l.J}$ rank lower than $y_{i.I}$ since $|i.J| = |l.J| = |i.I| - 1$. By induction hypothesis we can thus write

$$\hat{\lambda}(\delta_i y_I) = \hat{\delta}_j \hat{\delta}_i \hat{\lambda}(y_J) + \sum_{l=1}^{m} \hat{c}_{ijl}\hat{\delta}_l \hat{\lambda}(y_J).$$

As $\hat{\lambda}(y_J) = y_{\alpha-\varepsilon_j}$ we recognize that $\hat{\lambda}(\delta_i y_I) = \hat{\delta}_i \hat{\lambda}(y_I)$.

The case where I is not monotone is easily disposed off by induction since $y_I \equiv y_{\bar{I}} + p \bmod (\Omega)$ where p is a polynomial involving monotone derivatives ranking lower than $y_{\bar{I}}$. By the definition of ranking $y_{\bar{I}} \prec y_I$ and the conclusion follows easily. □

It follows that for all $p \in \mathbb{K}[\mathscr{T}_m(\mathscr{Y})]$ we have $\hat{\delta}_i \hat{\delta}_j(p) - \hat{\delta}_j \hat{\delta}_i(p) = \sum_{i=1}^{m} c_{ijl}\hat{\delta}_l(p)$.

Recall that for $I = (i_1, i_2, \ldots, i_p) \in \mathscr{W}_m$ we use the notation $\hat{\delta}^I = \hat{\delta}_{i_1}\hat{\delta}_{i_2}\ldots\hat{\delta}_{i_p}$. For $\alpha \in \mathscr{T}_m$ we take the convention that $\hat{\delta}^\alpha$ is $\hat{\delta}^{\overline{I}}$ where $\overline{I}$ is the only monotone word that corresponds to the term α, i.e. $\hat{\delta}^\alpha = \hat{\delta}_1^{\alpha_1}\ldots\hat{\delta}_m^{\alpha_m}$. Note that $\hat{\delta}^\alpha\hat{\delta}^\beta \neq \hat{\delta}^{\alpha+\beta}$.The derivation operators can nonetheless be normalized.

Proposition 5.2. *For all $I \in \mathscr{W}_m$ with $|I| \geqslant 2$ there is a family $\{a_L\}_{L\in\mathscr{M}_m}$ of polynomial functions in $\{\delta^K(c_{ijl}) \mid |K| \leqslant |I| - 2\}$ with coefficients in $\mathbb{Z}$ such that*

$$\hat{\delta}^I = \hat{\delta}^{\overline{I}} + \sum_{\substack{L\in\mathscr{M}_m\\ |L|<|I|}} a_L \hat{\delta}^L.$$

Proof. For all $I \in \mathscr{W}_m$ there is a permutation of the component σ_I that takes I to a monotone word. Any permutation can be written as the composition of a finite number of transpositions of neighboring elements. We call length the smallest number of such transpositions needed to decompose a permutation.

The proof is by induction on $|I|$ and the length of σ_I. The base cases are trivial: the result is true if $|I| = 2$ since $\hat{\delta}_i\hat{\delta}_j = \hat{\delta}_j\hat{\delta}_i + \sum_{l=1}^{m} c_{ijl}\hat{\delta}_l$ and if the length of σ_I is 0, meaning that I is monotone.

Take $I \in \mathscr{W}_m$, $|I| > 2$, and assume the result is true for all J s.t. either $|J| < |I|$ or $|J| = |I|$ but the length of σ_J is lower than the length of σ_I. We proceed to prove that the result is then true for I.

Take ρ to be the first transposition of neighboring elements in a minimal decomposition of σ_I and let J be the image of I by ρ. Then $\sigma_I = \sigma_J \circ \rho$ and the length of σ_J is strictly lower than the length of σ_I. There exist $L, K \in \mathscr{W}_m$ and $i, j \in \mathbb{N}_m$ s.t. $I = L.j.i.K$ and

$J = L.i.j.K$. Thus

$$\begin{aligned}\hat{\delta}^I &= \hat{\delta}^L \left(\hat{\delta}_i \hat{\delta}_j + \sum_{l=1}^{m} c_{jil} \hat{\delta}_l \right) \hat{\delta}^K \\ &= \hat{\delta}^J + \sum_{l=1}^{m} \sum_{M \star N = L} \hat{\delta}^M (c_{ijl}) \hat{\delta}^N \hat{\delta}_l \hat{\delta}^K,\end{aligned}$$

where $M \star N = L$ means that M is a subword of L and N is its complement in L.

On the one hand, $|N.l.K| < |I|$ so that we can apply the induction hypothesis to $\hat{\delta}[N.l.K] = \hat{\delta}^N \hat{\delta}_l \hat{\delta}^K$. On the other hand, σ_J has a length strictly smaller than σ_I so that we can also apply the induction hypothesis on $\hat{\delta}^J$. As $|M| \leqslant |I| - 2$, the conclusion follows for $\hat{\delta}^I$. $\square$

It is then just a matter of notation to obtain the following corollary.

Corollary 5.3. *For all* $\alpha, \beta \in \mathcal{T}_m$ *there is a family* $\{a_\gamma\}_{\gamma \in \mathcal{T}_m}$ *of polynomial functions in* $\{\delta^\mu(c_{ijl}) \,|\, |\mu| \leqslant |\alpha + \beta| - 2\}$ *with coefficients in* $\mathbb{Z}$ *such that*

$$\delta^\alpha \delta^\beta = \delta^{\alpha+\beta} + \sum_{|\gamma| < |\alpha+\beta|} a_\gamma \, \delta^\gamma.$$

Using the algebraic independence of the derivatives $\mathcal{T}_m(\mathcal{Y})$ we can actually show that $\{\hat{\delta}^\alpha \mid \alpha \in \mathbb{N}^m\}$ are linearly independent over $\mathbb{K}[\mathcal{T}_m(\mathcal{Y})]$. Thus the $\mathbb{K}[\mathcal{T}_m(\mathcal{Y})]$-algebra generated by $\hat{\Delta} = \{\hat{\delta}_1, \ldots, \hat{\delta}_m\}$ is a *PBW extension*, a notion introduced in [4], of $\mathbb{K}[\mathcal{T}_m(\mathcal{Y})]$.

We can also recall the property that $\hat{\lambda}(y_I) - \lambda(y_I)$ ranks less than $\lambda(y_I)$ as follows.

Proposition 5.4. *For all* $y \in \mathcal{Y}$ *and* $\alpha, \beta \in \mathcal{T}_m$ *we have* $\hat{\delta}^\beta(y_\alpha) - y_{\alpha+\beta}$ *ranks lower than* $y_{\alpha+\beta}$.

5.2. Definition

From the previous section we see that we can eventually drop the use of words. We shall thus gradually drop the use of $\mathcal{T}_m$ and $\mathcal{W}_m$ to retrieve the convenient notations in classical differential algebra.

Let $\mathcal{Y} = \{y_1, \ldots, y_n\}$ be a set of differential indeterminates. We consider the polynomial ring in infinitely many variables $\Theta\mathcal{Y} = \{y_\alpha \mid y \in \mathcal{Y},\ \alpha \in \mathbb{N}^m\}$, called the derivatives, with coefficients in a field $\mathbb{K}$ of characteristic zero. We endow $\mathbb{K}[\Theta\mathcal{Y}]$ with a set of m derivations $\Delta = \{\delta_1, \ldots, \delta_m\}$ for which $\mathbb{K}$ is a field of constants and defined recursively on $\Theta\mathcal{Y}$ by

$$\delta_i(y_\alpha) = \begin{cases} y_{\alpha+\varepsilon_i} & \text{if } \alpha_1 = \cdots = \alpha_{i-1} = 0, \\ \delta_j \delta_i(y_{\alpha-\varepsilon_j}) + \sum_{l=1}^m c_{ijl} \delta_l(y_{\alpha-\varepsilon_j}) & \text{where } j < i \text{ is s.t. } \alpha_j > 0 \\ & \text{while } \alpha_1 = \cdots = \alpha_{j-1} = 0, \end{cases}$$

where the family $\{c_{ijl}\}_{i,j,l\in\mathbb{N}_m}$ of elements of $\mathbb{K}[\Theta\mathscr{Y}]$ is such that for all $i, j, k, l \in \mathbb{N}_m$

- $c_{ijl} = -c_{jil}$,
- $\sum_{\mu=1}^{m} c_{ij\mu}c_{\mu kl} + c_{jk\mu}c_{\mu il} + c_{ki\mu}c_{\mu jl} = \delta_k(c_{ijl}) + \delta_i(c_{jkl}) + \delta_j(c_{kil})$.

For the sake of simplicity we introduce the following definition instead of the phrase *compatible semi-orderly ranking*.

Definition 5.5. An admissible ranking on $\mathbb{K}[\Theta\mathscr{Y}]$ is a total order $\prec$ on $\Theta\mathscr{Y}$ s.t.

- $|\alpha| < |\beta| \Rightarrow y_\alpha \prec y_\beta, \forall\alpha, \beta \in \mathbb{N}^m, \forall y \in \mathscr{Y}$,
- $y_\alpha \prec z_\beta \Rightarrow y_{\alpha+\gamma} \prec z_{\beta+\gamma}, \forall\alpha, \beta, \gamma \in \mathbb{N}^m, \forall y, z \in \mathscr{Y}$.
- $\sum_{l\in\mathbb{N}_m} c_{ijl}\delta_l(y_\alpha)$ ranks lower than $y_{\alpha+\varepsilon_i+\varepsilon_j}$ for all $y_\alpha \in \Theta\mathscr{Y}$ and all $i, j \in \mathbb{N}_m$.

We assume from now on that $\mathbb{K}[\Theta\mathscr{Y}]$ is endowed with an admissible ranking. By the previous section, we have

- $\delta_i\delta_j(p) - \delta_j\delta_i(p) = \sum_{l=1}^{m} c_{ijl}\delta_l(p), \quad \forall p \in \mathbb{K}[\Theta\mathscr{Y}]$,
- $\delta^\beta(y_\alpha) - y_{\alpha+\beta}$ ranks lower than $y_{\alpha+\beta}$.

In this case we shall say that $\mathbb{K}[\Theta\mathscr{Y}]$ is a *differential polynomial ring with nontrivial commutation rules for the derivations* $\Delta = \{\delta_1, \dots, \delta_m\}$ and write $\mathbb{K}[\![\mathscr{Y}]\!]$.

Let us make a couple of remarks. If the c_{ijl} are differential polynomials that involve derivative of order one or less than any orderly ranking is admissible. It is also possible to have the c_{ijl} involve derivatives of higher order.

Example 5.6. Assume $m = 1, n = 1$ and $c_{211} = y_{(2,0)}y_{(0,1)}$ and $c_{212} = -y_{(2,0)}y_{(1,0)}$. Then $c_{211}\delta_1(y) + c_{212}\delta_2(y) = 0$ and for all derivatives y_α with $|\alpha| \geqslant 1$ the property is clear.

Elimination rankings can also be admissible. If the coefficients c_{ijl} involve only derivatives (of order one or less) of a subset $\mathscr{Z} \subset \mathscr{Y}$ of the differential indeterminates then we can consider a ranking that eliminates $\mathscr{Y}\setminus\mathscr{Z}$. That is the case in the example of Section 2.

In classical differential algebra, i.e. when the derivation commute, rankings are only subject to the conditions

- $y_\alpha \prec y_{\alpha+\gamma}, \forall y \in \mathscr{Y},\ \alpha, \gamma \in \mathbb{N}^m$,
- $y_\alpha \prec z_\beta \Rightarrow y_{\alpha+\gamma} \prec z_{\beta+\gamma}, \forall\alpha, \beta, \gamma \in \mathbb{N}^m, \forall y, z \in \mathscr{Y}$.

There we can consider ranking that are not semi-orderly as the one given by

$$y_\alpha \prec y_\beta \Leftrightarrow\ \exists i \text{ s.t. } \alpha_1 = \beta_1 \text{ and } \dots \alpha_{i-1} = \beta_{i-1} \text{ and } \alpha_i < \beta_i.$$

The semi-orderly condition came in the proof of Theorem 4.8 and reappears for instance in the finite test of coherence.

The conditions on the coefficients c_{ijl} are satisfied if the derivations $\delta_1, \dots, \delta_m$ are linear combinations of m commuting derivations and the c_{ijl} are taken accordingly. To make it formal, consider a ring $\mathscr{R}$ on which the pairwise commuting derivations $\partial_1, \ldots, \partial_m$ act. Take $A=(a_{kl})$ a $m \times m$ matrix with entries in $\mathscr{R}$ s.t. its determinant is invertible in $\mathscr{R}$. Assume there is a ring isomorphism $\phi : \mathbb{K}[\Theta\mathscr{Y}] \to \mathscr{R}$ s.t. $\phi(\delta_i(p)) = \sum_{l=1}^m a_{il}\partial_l(\phi(p))$, for all $1 \leqslant i \leqslant m$. We can write this condition $\phi_*\delta_i = \sum_{l=1}^m a_{il}\partial_l$. $\phi_*\delta_1, \ldots, \phi_*\delta_m$ are derivations on $\mathscr{R}$ and therefore their commutators satisfies a Jacobi identity. Enforcing $\phi_*(\delta_i\delta_j - \delta_j\delta_i - \sum_{l=1}^m c_{ijl}\delta_l) = 0$ we obtain

$$\phi(c_{ijl}) = \sum_{k,\mu=1}^{m} \alpha_{lk}(a_{i\mu}\partial_\mu(a_{jk}) - a_{j\mu}\partial_\mu(a_{ik})),$$

where $(\alpha_{lk})_{kl}$ is the inverse of the transpose of A. The family $\{c_{ijl}\}_{i,j,l}$ then satisfies the required conditions.

The related theory presented in [23] can be seen to evolve around such a morphism $\phi : \mathbb{K}[\Theta\mathscr{Y}] \to \mathscr{R}$ where $\mathscr{R}$ is a classical differential polynomial ring with commuting derivations.

5.3. *Differential ideals*

A *differential ideal* in $\mathbb{K}[\![\mathscr{Y}]\!]$ is an ideal that is stable under the action of $\Delta = \{\delta_1, \ldots, \delta_m\}$. We shall note $[\Phi]$ the *differential ideal generated* by a non empty subset Φ of $\mathbb{K}[\![\mathscr{Y}]\!]$. $[\Phi]$ is defined as the intersection of all differential ideals containing Φ.

Proposition 5.7. *Let Φ be a non empty subset of $\mathbb{K}[\![\mathscr{Y}]\!]$. The differential ideal $[\Phi]$ is the ideal generated by $\Theta\Phi = \{\delta^\alpha\phi \mid \phi \in \Phi, \alpha \in \mathbb{N}^m\}$.*

Proof. The ideal generated by $\{\delta^I\phi \mid I \in \mathscr{W}_m, \phi \in \Phi\}$ is a differential ideal. It is furthermore contained in all the differential ideals containing Φ. It is thus equal to $[\Phi]$. We can conclude by Proposition 5.2. $\square$

That means that any element p of $[\Phi]$ can be written $p = \sum_{\phi\in\Phi,\alpha\in\mathbb{N}^m} a_{\alpha,\phi}\delta^\alpha\phi$ where the $a_{\alpha,\phi}$ form a family of $\mathbb{K}[\![\mathscr{Y}]\!]$ with finite support.

A differential ideal is *radical* if whenever a positive power of an element belongs to the differential ideal the element itself belongs to the differential ideal. $[\![\Phi]\!]$, the radical differential ideal generated by Φ, is defined as the intersection of all the radical differentials containing Φ. As in the classical setting, one shows that

$$[\![\Phi]\!] = \{p \in \mathbb{K}[\![\mathscr{Y}]\!] \mid \exists k \in \mathbb{N}\setminus\{0\},\ p^k \in [\Phi]\}.$$

If H is a subset of $\mathbb{K}[\![\mathscr{Y}]\!]$ we denote by H^∞ the monoid of elements that divide a power product of elements of H. Let I be a differential ideal in $\mathbb{K}[\![\mathscr{Y}]\!]$. The saturation of I by H is defined as

$$I : H^\infty = \{p \in \mathbb{K}[\![\mathscr{Y}]\!] \mid \exists h \in H^\infty,\ hp \in I\}.$$

The following property shows that $I : H^\infty$ is a differential ideal.

Lemma 5.8. *Let $p, q \in \mathbb{K}[\![\mathscr{Y}]\!]$ and $\alpha \in \mathbb{N}^m$. We have*

$$p^{|\alpha|+1}\delta^\alpha(q) \equiv p^{|\alpha|}\delta^\alpha(pq) \mod (\delta^\gamma(pq) \,|\, |\gamma| < |\alpha|).$$

Proof. The proof is by induction on $|\alpha|$. The result is true if $|\alpha| = 0$. Take α with $|\alpha| > 0$ and assume the result is true for all β with $|\beta| < |\alpha|$. Consider $i \in \mathbb{N}_m$ the smallest index with $\alpha_i \neq 0$. Take $\beta = \alpha - \varepsilon_i$ so that $|\beta| = |\alpha| - 1$ and $\delta_i\delta^\beta = \delta^\alpha$.

By induction hypothesis $p^{|\alpha|}\delta^\beta(q) \equiv p^{|\beta|}\delta^\beta(pq) \mod (\delta^\gamma(pq) \,|\, |\gamma| < |\beta|)$. So, we have, on the one hand,

$$p\delta_i(p^{|\alpha|}\delta^\beta(q)) \equiv |\beta|\delta_i(p)p^{|\beta|}\delta^\beta(pq) + p^{|\alpha|}\delta^\alpha(pq) \\ \mod (\delta^\gamma(pq), \delta_i\delta^\gamma(pq) \,|\, |\gamma| < |\beta|)$$

and, on the other hand,

$$p\delta_i\left(p^{|\alpha|}\,\delta^\beta(q)\right) = |\alpha|\delta_i(p)p^{|\alpha|}\delta^\beta(q) + p^{|\alpha|+1}\delta^\alpha(q).$$

As the induction hypothesis implies $p^{|\alpha|}\delta^\beta(q) \in (\delta^\gamma(pq) \,|\, |\gamma| < |\alpha|)$ and Corollary 5.3 implies $\delta_i\delta^\gamma(pq) \equiv \delta^{\varepsilon_i+\gamma}(pq) \mod (\delta^\mu(pq) \,|\, |\mu| \leqslant |\gamma|)$ we can conclude that

$$p^{|\alpha|+1}\delta^\alpha(pq) \equiv p\delta_i(p^{|\alpha|}\delta^\beta(q)) \equiv p^{|\alpha|}\delta^\alpha(pq) \mod (\delta^\gamma(pq) \,|\, |\gamma| < |\alpha|). \qquad \square$$

6. Constructive differential algebra

We generalize the constructions of classical differential algebra that lead to the fundamental theorems and to effective algorithms. After exhibiting the fundamental basic properties, the definitions and proofs are not really different than in the classical case [17,21]. We prove some key facts to see the tricks of the trade but for details we refer to [21] for fundamental theoretical results and to [17] for recent developments and algorithms.

6.1. Differential reduction and triangular sets

We assume that an admissible ranking is given on $\mathbb{K}[\![\mathscr{Y}]\!]$. Proposition 5.4 induces the following property that is the basis of differential reduction.

Proposition 6.1. *Assume that for $p \in \mathbb{K}[\![\mathscr{Y}]\!]$ we have $\mathsf{lead}(p) = y_\alpha$. For any $\beta \in \mathbb{N}^m$ with $|\beta| > 0$, $\mathsf{rank}(\delta^\beta(p)) = y_{\alpha+\beta}$ and $\mathsf{init}(\delta^\beta(p)) = \mathsf{sep}(p)$.*

Let p, q be elements of $\mathbb{K}[\![\mathscr{Y}]\!]$, $q \notin \mathbb{K}$ and $y_\alpha = \mathsf{lead}(q)$. The differential polynomial p is *partially reduced* w.r.t. q if it involves no derivative of the type $y_{\alpha+\gamma}$ where $|\gamma| > 0$. It is *reduced* if additionally the degree in y_α of p is lower than the one of q.

Thanks to Proposition 6.1 there is an algorithm pd-red based on derivation and (sparse) pseudo-division that, given p, q as before, returns a differential polynomial that is partially reduced w.r.t. q with the property that

$$\mathsf{pd\text{-}red}(p, q) \equiv sp \mod [q],$$

where s is a power product of factors of $\mathsf{sep}(p)$. Similarly $\mathsf{d\text{-}red}$ computes a differential polynomial that is reduced w.r.t. q with the property that

$$\mathsf{d\text{-}red}(p, q) \equiv hp \mod [q],$$

where h is a power product of factors of $\mathsf{sep}(p)$ and $\mathsf{init}(p)$.

Indeed, if p is not partially reduced w.r.t. q there is a $y_{\alpha+\gamma}$, $|\gamma| > 0$, present in p. Take the highest ranking such derivative. As $\mathsf{rank}(\delta^\gamma q) = y_{\alpha+\gamma}$, the remainder of the (sparse) pseudo-division of p by $\delta^\gamma q$ w.r.t. $y_{\alpha+\gamma}$ involves no derivatives $y_{\alpha+\beta}$ with $y_{\alpha+\gamma} \preceq y_{\alpha+\beta}$. Proceeding inductively on the remainders we obtain a differential polynomial with the required properties for $\mathsf{pd\text{-}red}(p, q)$. Adding a pseudo-division by q we obtain the algorithm $\mathsf{d\text{-}red}$.

We can thus define (*weak*) *differential triangular sets*, *differential chains* and *autoreduced sets* just as in [17, Section 3.2]. They are finite. If A is a differential triangular set we note, respectively, I_A and S_A the set of the initials and separants of its elements. Also $H_A = S_A \cup I_A$.

We can without any difficulty translate differential reduction by a weak differential triangular set from [17, Algorithms 3.12 and 3.13] as we did for reduction by a single differential polynomial. For a differential polynomial $p \in \mathbb{K}[\![\mathscr{Y}]\!]$ we can thus compute $\mathsf{d\text{-}red}(p, A)$ a differential polynomial that is reduced w.r.t. all the elements of A so that

$$\exists h \in H_A^\infty \text{ s.t. } hp \equiv \mathsf{d\text{-}red}(p, A) \mod [A]$$

and similarly for $\mathsf{pd\text{-}red}$ with $h \in S_A^\infty$.

6.2. Coherence

The keystone of characteristic decomposition algorithms in the classical case is coherence and the related Rosenfeld's lemma [37]. We proceed to show the result in our new setting. There is no essential difference with the classical case. We shall use nonetheless the fact that the ranking is semi-orderly in the finite test for coherence.

Coherence and Rosenfeld's lemma provide an analogue to the S-polynomial criterion for Gröbner basis. The proof of Rosenfeld's lemma is close to the test for Gröbner basis in terms of t-presentation developed in the context of the Buchberger second criterion [3, Sections 5.4 and 5.5]. It is easier than an approach through a diamond lemma.

For $y_\alpha \in \Theta\mathscr{Y}$ we note $\Theta\mathscr{Y}_{<y_\alpha}$ the set of derivatives that rank lower than y_α. Let A be a differential triangular set. ΘA denotes the set $\{\delta^\alpha a \mid \alpha \in \mathbb{N}^m, a \in A\}$. We note $\Theta A_{<y_\alpha} = \Theta A \cap \Theta\mathscr{Y}_{<y_\alpha}$.

Definition 6.2. Let A be a d-triangular set in $\mathscr{F}[\![\mathscr{Y}]\!]$ and $H \supset S_A$ a subset of $\mathscr{F}[\![\mathscr{Y}]\!]$. A is said to be coherent away from H (or H-coherent for short) if whenever $a, b \in A$ are such that $\mathsf{lead}(a) = y_\alpha$ and $\mathsf{lead}(b) = y_\beta$ for some $y \in \mathscr{Y}$ and $\alpha, \beta \in \mathbb{N}^m$ then for any $\gamma \in (\alpha + \mathbb{N}^m) \cap (\beta + \mathbb{N}^m)$ we have

$$\mathsf{sep}(b)\, \delta^{\gamma-\alpha}(a) - \mathsf{sep}(a)\, \delta^{\gamma-\beta}(b) \in (\Theta A_{<y_\gamma}) : H^\infty.$$

Theorem 6.3. *Let A be a d-triangular set and $H \supset S_A$ a set of differential polynomials partially reduced w.r.t. A. If A is H-coherent then any differential polynomial of $[A] : H^\infty$ that is partially reduced w.r.t. A belongs to $(A) : H^\infty$.*

Proof. For $a \in A$ we note u_a and s_a, respectively, the leader and the separant of a. Let us consider $p \in [A] : H^\infty$. There thus exists a finite subset D of $\mathbb{N}^m \setminus \{0\} \times A$ s.t. for some $h \in H^\infty$ we can write

$$hp = \sum_{(\alpha,a)\in D} p_{\alpha,a}\delta^\alpha(a) + \sum_{a\in A} p_a a \tag{1}$$

for some $p_a, p_{\alpha,a} \in \mathbb{K}[\![\mathscr{Y}]\!]$. For each equation of type (1) we consider v to be the highest ranking derivative of $\mathsf{lead}(\Theta^+(A))$ that appears effectively on the right-hand side.

Assume that p is partially reduced w.r.t. A. If the set D is empty then $p \in (A) : H^\infty$. Assume, for contradiction, that there is no relation of type (1) with an empty D for p. Among all the possible relationships (1) that can be written, we consider one for which v is minimal.

Consider $E = \{(\alpha, a) \in D \mid \mathsf{lead}(\delta^\alpha(a)) = v\}$ and single out any $(\bar\alpha, \bar a)$ of E. As A is H-coherent, for all (α, a) of E we have $s_{\bar a}\delta^\alpha(a) \equiv s_a\delta^{\bar\alpha}(\bar a) \bmod (\Theta A_{<v}) : H^\infty$. Thus

$$\begin{aligned} s_{\bar a}hp \equiv & \left(\sum_{(\alpha,a)\in E} s_a p_{\alpha,a}\right)\delta^{\bar\alpha}(\bar a) + \sum_{(\alpha,a)\in D\setminus E} s_{\bar a} p_{\alpha,a}\delta^\alpha a \\ & + \sum_{a\in A} s_{\bar a} p_a a \ \bmod \ (\Theta A_{<v}) : H^\infty \end{aligned} \tag{2}$$

so that we can find $k \in H^\infty$ s.t.

$$kp = q_{\bar\alpha,\bar a}\delta^{\bar\alpha}(\bar a) + \sum_{\substack{(\alpha,a)\in\mathbb{N}^m\setminus\{0\}\times A,\\ \mathsf{lead}(\delta^\alpha(a))<v}} q_{\alpha,a}\,\delta^\alpha a + \sum_{a\in A} q_a a \tag{3}$$

for some $q_a, q_{\alpha,a} \in \mathbb{K}[\![\mathscr{Y}]\!]$.

We proceed now to eliminate v from the coefficients $q_{\bar\alpha,\bar a}$ and $q_{\alpha,a}$ and q_a. We make use of the fact that $s_{\bar a}\, v = \delta^{\bar\alpha}(\bar a) - \mathsf{tail}(\delta^{\bar\alpha}(\bar a))$. Recall that $\mathsf{tail}(\delta^{\bar\alpha}(\bar a))$ contains only derivatives lower than v. Multiplying both sides of (3) by $s_{\bar a}^d$, where d is the degree of v in the right-hand side, and replacing $s_{\bar a}v$ by $\delta^{\bar\alpha}(\bar a) - \mathsf{tail}(\delta^{\bar\alpha}(\bar a))$ we can rewrite the relationship obtained as

$$s_{\bar a}^d kp = r_d\delta^{\bar\alpha}(\bar a)^d + \cdots + r_1\delta^{\bar\alpha}(\bar a) + r_0, \tag{4}$$

where $r_0, r_1, \ldots, r_d$ no longer contain v and $r_0 \in (\Theta A_{<v})$. The only occurrences of v in that right-hand side is through $\delta^{\bar\alpha}(\bar a)$. Because p and the elements of H are partially reduced w.r.t. A, v does not appear in the left-hand side. The coefficients r_i, for $1 \leqslant i \leqslant d$ must be zero. We have thus exhibited a relationship like (1) with a v lower than what we started from. This contradicts our hypotheses. □

We proceed to give a finite test for coherence. Note before that derivations of products are given by the usual formula.

Proposition 6.4. *Let $p, q \in \mathbb{K}[\![\mathscr{Y}]\!]$ and $\alpha \in \mathbb{N}^m$. We have $\delta^\alpha(pq) = \sum_{\mu+\nu=\alpha} \delta^\mu(p)\delta^\nu(q)$.*

Lemma 6.5. *Let A be a d-triangular set and H a finite subset of $\mathbb{K}[\![\mathscr{Y}]\!]$ and $y_\alpha \in \Theta\mathscr{Y}$. If $p \in (\Theta A_{<y_\alpha}) : H^\infty$ then $\delta^\beta(p) \in (\Theta A_{<y_{\alpha+\beta}}) : H^\infty$.*

Proof. From Propositions 6.4 and 6.1 one easily deduces that if $q \in (\Theta A_{<y_\alpha})$ then $\delta^\beta(q) \in (\Theta A_{<y_{\alpha+\beta}})$.

As $p \in (\Theta A_{<y_\alpha}) : H^\infty$ there exists $h \in H^\infty$ s.t. $hp \in (\Theta A_{<y_\alpha})$. By Lemma 5.8

$$h^{|\beta|+1}\delta^\beta(p) \equiv h^{|\beta|}\delta^\beta(hp) \bmod (\delta^\gamma(hp) \mid |\gamma| < |\beta|).$$

As the ranking is semi-orderly $|\gamma| < |\beta|$ implies that $y_{\alpha+\gamma} \prec y_{\alpha+\beta}$. So by induction on $|\beta|$ we prove that $h^{|\beta|+1}\delta^\beta(p) \in (\Theta A_{<y_{\alpha+\beta}})$. The conclusion follows. □

For $\alpha, \beta \in \mathbb{N}^m$, we denote $\alpha \diamond \beta$ the element of $\mathbb{N}^m$ having $\max(\alpha_i, \beta_i)$ for ith component. Any element of $(\alpha + \mathbb{N}^m) \cap (\beta + \mathbb{N}^m)$ can be written $\alpha \diamond \beta + \mu$ for some $\mu \in \mathbb{N}^m$.

Proposition 6.6. *Let A be a weak d-triangular set and $H \supset S_A$ a subset of $\mathbb{K}[\![\mathscr{Y}]\!]$. If for all $a, b \in A$ s.t. $\mathsf{lead}(a) = y_\alpha$ and $\mathsf{lead}(b) = y_\beta$ for some $y \in \mathscr{Y}$ and $\alpha, \beta \in \mathbb{N}^m$ we have*

$$\Delta(a, b) = \mathsf{sep}(b)\delta^{\alpha\diamond\beta-\alpha}(a) - \mathsf{sep}(a)\delta^{\alpha\diamond\beta-\beta}(b) \in (\Theta A_{<y_{\alpha\diamond\beta}}) : H^\infty$$

then A is H-coherent.

Proof. We deduce from Proposition 6.4 that for $h, p \in \mathbb{K}[\![\mathscr{Y}]\!]$ and $\gamma \in \mathbb{N}^m$, $\delta^\gamma(hp) \equiv h\delta^\gamma(p) \bmod (\delta^\mu(p) \mid |\mu| < |\gamma|)$. Take $\gamma = \alpha \diamond \beta$ and $\gamma + \nu$ an element of $(\alpha + \mathbb{N}^m) \cap (\beta + \mathbb{N}^m)$. Then $\Delta(a, b) = \mathsf{sep}(b)\delta^{\gamma-\alpha}(a) - \mathsf{sep}(a)\delta^{\gamma-\beta}(b)$ and from the first remark

$$\begin{aligned}\delta^\mu(\Delta(a, b)) \equiv{}& \mathsf{sep}(b)\delta^\mu\delta^{\gamma-\alpha}(a) - \mathsf{sep}(a)\delta^\mu\delta^{\gamma-\beta}(b)\\ &\bmod (\delta^\nu\delta^{\gamma-\alpha}(a), \delta^\nu\delta^{\gamma-\beta}(b) \mid |\nu| < |\mu|).\end{aligned}$$

According to the hypotheses and Lemma 6.5, $\delta^\mu(\Delta(a, b))$ belongs to $(\Theta A_{<y_{\gamma+\mu}}) : H^\infty$. Now $\delta^\mu\delta^{\gamma-\alpha}(a) \equiv \delta^{\mu+\gamma-\alpha}(a) \bmod (\delta^\nu(a) \mid |\nu| < |\mu + \gamma - \alpha|)$ by Corollary 5.3 and $\mathsf{lead}(\delta^\nu(a)) = y_{\alpha+\nu} < y_{\gamma+\mu}$ for $|\nu| < |\mu + \gamma - \alpha|$ by Proposition 6.1 and the fact that the ranking is semi-orderly. Similarly for $\delta^\mu\delta^{\gamma-\beta}(b)$. The conclusion follows. □

6.3. Characteristic decomposition

The definitions and first properties of differential characteristic sets, characterizable differential ideals and characteristic decomposition are exposed in details in [17, Sections 3.3 and 5]. This is based on the grounds of differential reduction and we saw that it goes through without difficulty. We just sketch the content of those sections in our new setting.

We define *characteristic sets* of differential ideals as differential chains of minimal rank in those differential ideals. Any differential ideal admits a characteristic set. If C is a characteristic set of a differential ideal I then $C \subset I \subset [C] : H_C^\infty$ and $p \in I \Rightarrow \mathsf{d\text{-}red}(p, C) = 0$.

We can define as in [15,17] *characterizable* differential ideals. They are the differential ideals $[C] : H_C^\infty$ where C is a characteristic set of $[C] : H_C^\infty$. Consequently $p \in [C] : H_C^\infty \Leftrightarrow \mathsf{d\text{-}red}(p, C) = 0$. As expanded upon in [17, Section 5.1] C is a characteristic set of $[C] : H_C^\infty$ iff C is a *regular differential chain*. In that case $[C] : H_C^\infty = [C] : S_C^\infty$. Some radical differential ideals can be characterizable for some ranking but not for another one.

Note that prime differential ideals are characterizable for any admissible ranking. Therefore, a characteristic set of a prime ideal gives a finite representation for this prime differential ideal. This finiteness property is at the heart of the basis theorem [21, Chapter III] that asserts that any radical differential ideal is finitely generated (as a radical differential ideal). We revisit the proof of this theorem in the appendix. It goes along with the proof that any radical differential ideal is the intersection of finitely many prime ideals.

Rosenfeld's lemma is the keystone of the characteristic decomposition algorithms [7,8,21] in the classical case. For the algorithms of [15,17] we additionally need the fact that any radical differential ideal is the intersection of finitely many prime ideals. As we proved a version of Rosenfeld's lemma in our new setting, all the algorithms go through. We thus only state the existence of those algorithms in our new setting.

Assume $\mathbb{K}[\![\mathscr{Y}]\!]$ is endowed with an admissible ranking. Consider a finite set Φ of differential polynomials in $\mathbb{K}[\![\mathscr{Y}]\!]$. Algorithms [7,8,15,17,21] are easily translated to compute a finite set $\mathscr{C}$ of regular differential chains such that

$$[\![\Phi]\!] = \bigcap_{C \in \mathscr{C}} [C] : S_C^\infty.$$

The characteristic decomposition algorithm allow to test membership and answer differential elimination questions. For examples of use of characteristic decomposition, in the classical case, see [17, Section 8].

Acknowledgements

E. Mansfield is the one who dragged me into this project and it has been such an exciting experience. I wish to express here my deep gratitude to her. She and I. Kogan spent precious time explaining to me the moving frame constructions. It has been fun and I want to thank them both for that. Discussion with I. Kogan led to the reduction for which I am developing this theory and the material in this paper took a definite shape while I visited M. Singer at MSRI. Discussions with him were absolutely decisive in finding the right way to look at the problem. I would like to thank MSRI for giving me this opportunity and providing support which made this possible.

Appendix A. The basis theorem

We adapt here the proof of the basis theorem as disseminated over several chapters in [21]. Beside rewriting it in our new setting we make the simplifications owing to the fact that we are dealing with characteristic zero only. We also give the result about the decomposition of radical differential ideals into prime ideals that is an easy consequence of the basis theorem.

Proposition A.1. *For all subset $\Phi, \Psi \subset \mathbb{K}[\![\mathcal{Y}]\!]$ we have $[\![\Phi]\!] \cap [\![\Psi]\!] = [\![\Psi \cdot \Phi]\!]$, where $\Psi \cdot \Phi$ denotes the set of all the products of the pairs in $\Psi \times \Phi$.*

Proof. Trivially $\Psi \cdot \Phi$ is a subset of $[\![\Phi]\!]$ and $[\![\Psi]\!]$ and therefore of their intersection. Conversely, take $p \in [\![\Phi]\!] \cap [\![\Psi]\!]$. There exist $r, s \in \mathbb{N} \backslash \{0\}$ s.t.

$$p^r = \sum_{\phi \in \Phi, \alpha \in \mathbb{N}^m} a_{\phi,\alpha} \delta^\alpha(\phi),$$
$$p^s = \sum_{\psi \in \Psi, \beta \in \mathbb{N}^m} b_{\psi,\beta} \delta^\beta(\psi)$$

so that

$$p^{r+s} = \sum_{\phi \in \Phi, \psi \in \Psi, \alpha, \beta \in \mathbb{N}^m} a_{\phi,\alpha} b_{\psi,\beta} \delta^\alpha(\phi)\, \delta^\beta(\psi).$$

From Lemma 5.8 we easily deduce that $\delta^\alpha(\phi)\delta^\beta(\psi) \in [\![\phi\, \psi]\!]$ so that $p \in [\![\Psi \cdot \Phi]\!]$. □

Lemma A.2. *Let $\mathcal{F}$ be the set of radical differential ideals in $\mathbb{K}[\![\mathcal{Y}]\!]$ that are not finitely generated. $\mathcal{F}$ has a maximal element* (*for the partial order of inclusion*) *and all maximal elements are prime.*

Proof. The first part is by Zorn's lemma that says: if every chain in $\mathcal{F}$ has an upper bound then $\mathcal{F}$ has a maximal element. Consider $\{J_i\}_{i \in \mathbb{N}}$ a family in $\mathcal{F}$ forming a chain, i.e. s.t. $J_i \subsetneq J_{i+1}$. $J = \bigcup_{i=1}^{\infty} J_i$ is also a radical differential ideal. If there existed a finite subset Φ s.t. $J = [\![\Phi]\!]$ then Φ would be contained in some J_k for some $k \in \mathbb{N}$ and we would have $J_k = J_{k+1} = \cdots = J = [\![\Phi]\!]$. Thus J belongs to $\mathcal{F}$.

Assume now that P is a maximal element of $\mathcal{F}$ and take $a, b \notin P$. As P is maximal in $\mathcal{F}$, there exist Φ, Ψ finite subsets of $\mathbb{K}[\![\mathcal{Y}]\!]$ s.t. $[\![P \cup \{a\}]\!] = [\![\Phi]\!]$ and $[\![P \cup \{b\}]\!] = [\![\Psi]\!]$. As

$$[\![P \cup \{ab\}]\!] = [\![P \cup \{a\}]\!] \cap [\![P \cup \{b\}]\!] = [\![\Phi]\!] \cap [\![\Psi]\!] = [\![\Phi \cdot \Psi]\!]$$

ab cannot belong to P since otherwise it would contradict the hypothesis that P is not finitely generated. Thus P is prime. □

Theorem A.3 (*The basis theorem*). *For any radical differential ideal J in $\mathbb{K}[\![\mathcal{Y}]\!]$ there exists a finite subset Φ of $\mathbb{K}[\![\mathcal{Y}]\!]$ such that $J = [\![\Phi]\!]$.*

Proof. Assume for contradiction that the set $\mathcal{F}$ of radical differential ideals that are not finitely generated is non empty. By Lemma A.2, we can consider P a maximal element in $\mathcal{F}$. P is prime. Let C be a characteristic set for P and h be the product of the initials and separants of C. Then $P = [C] : H_C^\infty = [\![C]\!] : h$.

As $h \notin P$ and P is maximal in $\mathcal{F}$ there exists a finite subset Φ of $\mathbb{K}[\![\mathcal{Y}]\!]$ s.t. $[\![P \cup \{h\}]\!] = [\![\Phi]\!]$. Any element of Φ can be written as a linear combination of h and its derivatives and elements of P. Let Ψ be the finite set of elements of P coming into those linear combinations. Then $[\![P \cup \{h\}]\!] = [\![\Phi]\!] = [\![\{h\} \cup \Psi]\!]$.

As $P = P : h \cap [\![P \cup \{h\}]\!] = P \cap [\![\{h\} \cup \Psi]\!]$, by Corollary A.1, $P = [\![P \cdot (\{h\} \cup \Psi)]\!]$ and as $P \cdot \Psi \subset [\![\Psi]\!]$ we have $P = [\![(P \cdot \{h\}) \cup \Psi]\!]$. Now $P = [\![C]\!] : h$ so that $P \cdot \{h\} \subset [\![C]\!] \subset P$. Thus $P \subset [\![C \cup \Psi]\!] \subset P$. It follows that $P = [\![C \cup \Psi]\!]$ is finitely generated. $\square$

Theorem A.4. *In $\mathbb{K}[\![\mathscr{Y}]\!]$ any radical differential ideal is the intersection of a finite number of prime differential ideals. The set of prime differential ideals coming into such a decomposition with no superfluous component is unique.*

Proof. As any radical differential ideal of $\mathbb{K}[\![\mathscr{Y}]\!]$ is finitely generated any ascending chain of radical differential ideal is finite. Thus among any set of radical differential ideal there is one that is maximal w.r.t. inclusion.

Let J be a radical differential ideal and p be an element that does not belong to J. We show that there is a prime differential ideal that contains J and not p. Take P to be a radical differential ideal that contains J and not p that is maximal w.r.t. that property. Take $a, b \in \mathbb{K}[\![\mathscr{Y}]\!]$ s.t. $ab \in P$ so that $P = [\![P \cup \{a\}]\!] \cap [\![P \cup \{b\}]\!]$. If neither a nor b belonged to P then p would belong to $[\![P \cup \{a\}]\!]$ and $[\![P \cup \{b\}]\!]$ by the maximality hypothesis on P. This cannot be the case since p does not belong to their intersection. P is prime. Thus J is the intersection of all the prime differential ideals that contain it.

Assume the set of radical differential ideals that cannot be written as a finite intersection of prime differential ideal is not empty and consider J a maximal element in that set. Obviously J is not prime. Thus, there exists $a, b \notin J$ s.t. $ab \in J$. Then $J = [\![J \cup \{a\}]\!] \cap [\![J \cup \{b\}]\!]$. Since $J \subsetneq [\![J \cup \{a\}]\!]$ it must be that $[\![J \cup \{a\}]\!]$ is an intersection of a finite number of prime differential ideals and similarly for $[\![J \cup \{b\}]\!]$. Thus, J has to be an intersection of a finite number of prime differential ideals.

Assume that $J = \bigcap_{P \in \mathscr{P}} P = \bigcap_{Q \in \mathscr{Q}} Q$ where $\mathscr{P}$ and $\mathscr{Q}$ are finite set of prime differential ideals such that no element of it contains another one. Then for any $P \in \mathscr{P}$, $\bigcap_{Q \in \mathscr{Q}} Q \subset P$ so that there exists Q such that $Q \subset P$. Similarly there is a $P' \in \mathscr{P}$ s.t. $P' \subset Q$. It must be that $P = Q = P'$. Thus $\mathscr{P} = \mathscr{Q}$. $\square$

References

[1] I. Anderson, Vessiot, a MAPLE package for differential geometry, Formal Geometry and Mathematical Physics Group, http://www.math.usu.edu/fg_mp/Pages/SymbolicsPage/Symbolics.html

[2] J. Apel, Computational ideal theory in finitely generated extension rings, Theoret. Comput. Sci. 244 (1–2) (2000) 1–33.

[3] T. Becker, V. Weispfenning, Gröbner Bases—A Computational Approach to Commutative Algebra, Springer, New York, 1993.

[4] A.D. Bell, K.R. Goodearl, Uniform rank over differential operator rings and Poincaré–Birkhoff–Witt extensions, Pacific J. Math. 131 (1) (1988) 13–37.

[5] F. Boulier, E. Hubert, DIFFALG: description, help pages and examples of use, Symbolic Computation Group, University of Waterloo, Ontario, Canada, 1998, now at http://www.inria.fr/cafe/Evelyne.Hubert/diffalg

[6] F. Boulier, D. Lazard, F. Ollivier, M. Petitot, Representation for the radical of a finitely generated differential ideal, in: A.H.M. Levelt (Ed.), ISSAC'95, ACM Press, New York, 1995.

[7] F. Boulier, D. Lazard, F. Ollivier, M. Petitot, Computing representations for radicals of finitely generated differential ideals, Technical Report IT-306, LIFL, 1997.

[8] D. Bouziane, A. Kandri Rody, H. Maârouf, Unmixed-dimensional decomposition of a finitely generated perfect differential ideal, J. Symbolic Comput. 31 (6) (2001) 631–649.
[9] J. Carminati, K. Vu, Symbolic computation and differential equations: Lie symmetries, J. Symbolic Comput. 29 (1) (2000) 95–116.
[10] G. Carra Ferro, Gröbner bases and differential algebra, in: AAECC, Lecture Notes in Computer Science, vol. 356, Springer, New York, 1987.
[11] P.J. Cassidy, Differential algebraic groups, Amer. J. Math. 94 (1972) 891–954.
[12] J. Douglas, Solution of the inverse problem of the calculus of variations, Trans. Amer. Math. Soc. 50 (1941) 71–128.
[13] M. Fels, P.J. Olver, Moving coframes. II. Regularization and theoretical foundations, Acta Appl. Math. 55 (2) (1999) 127–208.
[14] M.W. Giesbrecht, G.J. Reid, Y. Zhang, Non-commutative Gröbner bases in Poincaré–Birkhoff–Witt extensions, in: V.G. Ganzha, E.W. Mayr, E.V. Vorozhtsov (Eds.), Proceedings of the Fifth International Workshop on Computer Algebra in Scientific Computing, CASC 2002, Yalta, Ukraine, Technical University of Munich, 2002, pp. 97–106.
[15] E. Hubert, Factorisation free decomposition algorithms in differential algebra, J. Symbolic Comput. 29 (4–5) (2000) 641–662.
[16] E. Hubert, Notes on triangular sets and triangulation-decomposition algorithms I: polynomial systems, in: F. Winkler, U. Langer (Eds.), Symbolic and Numerical Scientific Computing 2001, Lecture Notes in Computer Science, vol. 2630, Springer, Heidelberg, 2003, pp. 1–39.
[17] E. Hubert, Notes on triangular sets and triangulation-decomposition algorithms II: differential systems, in: F. Winkler, U. Langer (Eds.), Symbolic and Numerical Scientific Computing 2001, Lecture Notes in Computer Science, vol. 2630, Springer, Heidelberg, 2003, pp. 40–87.
[18] E. Hubert, I. Kogan, E. Mansfield, Algebraic moving frame construction, in preparation.
[19] A. Kandri-Rody, V. Weispfenning, Noncommutative Gröbner bases in algebras of solvable type, J. Symbolic Comput. 9 (1) (1990) 1–26.
[20] I.A. Kogan, P.J. Olver, Invariant Euler–Lagrange equations and the invariant variational bicomplex, Acta Appl. Math. 76 (2) (2003) 137–193.
[21] E.R. Kolchin, Differential Algebra and Algebraic Groups, Pure and Applied Mathematics, vol. 54, Academic Press, New York, 1973.
[22] E.R. Kolchin, Differential Algebraic Groups, Pure and Applied Mathematics, vol. 114, Academic Press, Orlando, FL, 1985.
[23] F. Lemaire, G. Reid, Y. Zhang, Non-commutative existence and uniqueness theory for analytic systems of nonlinear partial differential equations, `http://www.orcca.on.ca/~reid/`
[24] V. Levandovskyy, H. Schönemann, PLURAL—a computer algebra system for noncommutative polynomial algebras, in: Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2003, pp. 176–183.
[25] I.G. Lisle, Equivalence transformation for classes of differential equations, Ph.D. Thesis, University of British Columbia, 1992.
[26] I.G. Lisle, G.J. Reid, Symmetry classification using invariant moving frame, `http://www.cecm.sfu.ca/~reid/`
[27] B. Malgrange, Systèmes différentiels involutifs, SMF, to appear.
[28] E.L. Mansfield, Differential Gröbner bases, Ph.D. Thesis, University of Sydney, 1991.
[29] E.L. Mansfield, The differential algebra package `diffgrob2`, MapleTech, vol. 3, 1996, pp. 33–37.
[30] E.L. Mansfield, Algorithms for symmetric differential systems, Found. Comput. Math. 1 (4) (2001) 335–383.
[31] E.L. Mansfield, A. Szanto, Elimination theory for differential difference polynomials, in: Proceedings of the 2003 International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2003, pp. 191–198.
[32] T. Mora, An introduction to commutative and noncommutative Gröbner bases, Theoret. Comput. Sci. 134 (1) (1994) 131–173 Second International Colloquium on Words, Languages and Combinatorics, Kyoto, 1992.
[33] S. Neut, Implantation et nouvelles applications de la méthode d'équivalence de Cartan, Ph.D. Thesis, Université des Sciences et Technologie de Lille, `http://www.lifl.fr/~neut`, 2003.
[34] F. Ollivier, Standard bases of differential ideals, in: Applied Algebra, Algebraic Algorithms and Error-correcting Codes, (Tokyo, 1990), Springer, Berlin, 1991, pp. 304–321.

[35] G.J. Reid, A.D. Wittkopf, A. Boulton, Reduction of systems of nonlinear partial differential equations to simplified involutive forms, European J. Appl. Math. 7 (1996) 604–635.
[36] J.F. Ritt, Differential Algebra, Colloquium Publications, vol. XXXIII, American Mathematical Society, Providence, RI, 1950, http://www.ams.org/online_bks
[37] A. Rosenfeld, Specializations in differential algebra, Trans. Amer. Math. Soc. 90 (1959) 394–407.
[38] C. Rust, Rankings of derivatives for elimination algorithms and formal solvability of analytic partial differential equations, Ph.D. Thesis, University of Chicago, 1998.
[39] A. Witkopf, G. Reid, The RIF package, CECM, Simon Fraser University, Vancouver, http://www.cecm.sfu.ca/wittkopf/rif.html
[40] Y. Yaffe, Model completion of Lie differential fields, Ann. Pure Appl. Logic 107 (1–3) (2001) 49–86.